

**IJCSIS Vol. 12 No. 9, September 2014**  
**ISSN 1947-5500**

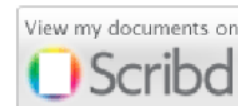
# **International Journal of Computer Science & Information Security**

**© IJCSIS PUBLICATION 2014**



Cogprints

Google scholar



SciRate.com

CiteSeer<sup>x</sup> beta



# IJCSIS

ISSN (online): 1947-5500

Please consider to contribute to and/or forward to the appropriate groups the following opportunity to submit and publish original scientific results.

## CALL FOR PAPERS

### International Journal of Computer Science and Information Security (IJCSIS) January-December 2014 Issues

The topics suggested by this issue can be discussed in term of concepts, surveys, state of the art, research, standards, implementations, running experiments, applications, and industrial case studies. Authors are invited to submit complete unpublished papers, which are not under review in any other conference or journal in the following, but not limited to, topic areas.

See authors guide for manuscript preparation and submission guidelines.

**Indexed by Google Scholar, DBLP, CiteSeerX, Directory for Open Access Journal (DOAJ), Bielefeld Academic Search Engine (BASE), SCIRUS, Scopus Database, Cornell University Library, ScientificCommons, ProQuest, EBSCO and more.**

**Deadline:** see web site

**Notification:** see web site

**Revision:** see web site

**Publication:** see web site

Context-aware systems  
Networking technologies  
Security in network, systems, and applications  
Evolutionary computation  
Industrial systems  
Evolutionary computation  
Autonomic and autonomous systems  
Bio-technologies  
Knowledge data systems  
Mobile and distance education  
Intelligent techniques, logics and systems  
Knowledge processing  
Information technologies  
Internet and web technologies  
Digital information processing  
Cognitive science and knowledge

Agent-based systems  
Mobility and multimedia systems  
Systems performance  
Networking and telecommunications  
Software development and deployment  
Knowledge virtualization  
Systems and networks on the chip  
Knowledge for global defense  
Information Systems [IS]  
IPv6 Today - Technology and deployment  
Modeling  
Software Engineering  
Optimization  
Complexity  
Natural Language Processing  
Speech Synthesis  
Data Mining

For more topics, please see web site <https://sites.google.com/site/ijcsis/>

arXiv.org Google scholar

SCIRUS  
search engine for science

ScientificCommons

Scribd

docstoc  
find and share professional documents

BASE  
Bielefeld Academic Search Engine

CiteSeer<sup>x</sup> beta

dblp.uni-trier.de  
Computer Science  
Bibliography

DOAJ  
DIRECTORY OF  
OPEN ACCESS  
JOURNALS



ProQuest

For more information, please visit the journal website (<https://sites.google.com/site/ijcsis/>)

## Editorial

### Message from Managing Editor

*The **International Journal of Computer Science and Information Security (IJCSIS)** promotes research work publications which offer a significant contribution to the computer science knowledge, and which are of high interest to a wide academic/research/practitioner audience. Coverage extends to all main-stream and state of the art branches of computer science, security and related information technology applications. As a scholarly open access peer-reviewed journal, IJCSIS mission is to provide an outlet for quality research articles. It aims to promote universal access with equal opportunities for international scientific community; to scientific knowledge, and the creation, and dissemination of scientific and technical information.*

*IJCSIS archives all publications in major academic/scientific databases. Indexed by the following International agencies and institutions: Google Scholar, Bielefeld Academic Search Engine (BASE), CiteSeerX, SCIRUS, Cornell's University Library EI, Scopus, DBLP, DOI, ProQuest, EBSCO. Google Scholar reported increased in number cited papers published in IJCSIS (**No. of Cited Papers:524, No. of Citations:1043, Years:5**). Abstracting/indexing, editorial board and other important information are available online on homepage. This journal supports the Open Access policy of distribution of published manuscripts, ensuring "free availability on the public Internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of [published] articles".*

*IJCSIS editorial board, consisting of international experts, ensures a rigorous peer-reviewing process. We look forward to your collaboration. For further questions please do not hesitate to contact us at [ijcsiseditor@gmail.com](mailto:ijcsiseditor@gmail.com).*

*A complete list of journals can be found at:*

*<http://sites.google.com/site/ijcsis/>*

*IJCSIS Vol. 12, No. 9, September 2014 Edition*

*ISSN 1947-5500 © IJCSIS, USA.*

*Journal Indexed by (among others):*



## IJCSIS EDITORIAL BOARD

**Dr. Yong Li**

School of Electronic and Information Engineering, Beijing Jiaotong University,  
P. R. China

**Prof. Hamid Reza Naji**

Department of Computer Engineering, Shahid Beheshti University, Tehran, Iran

**Dr. Sanjay Jasola**

Professor and Dean, School of Information and Communication Technology,  
Gautam Buddha University

**Dr Riktesh Srivastava**

Assistant Professor, Information Systems, Skyline University College, University  
City of Sharjah, Sharjah, PO 1797, UAE

**Dr. Siddhivinayak Kulkarni**

University of Ballarat, Ballarat, Victoria, Australia

**Professor (Dr) Mokhtar Beldjehem**

Sainte-Anne University, Halifax, NS, Canada

**Dr. Alex Pappachen James (Research Fellow)**

Queensland Micro-nanotechnology center, Griffith University, Australia

**Dr. T. C. Manjunath**

HKBK College of Engg., Bangalore, India.

**Prof. Elboukhari Mohamed**

Department of Computer Science,  
University Mohammed First, Oujda, Morocco

# TABLE OF CONTENTS

## **1. Paper 31081412: An Improved UGS Scheduling with QoE Metrics in WiMAX Network (pp. 1-6)**

*Tarik ANOUARI (1), Abdelkrim HAQIQ (1, 2)*

*(1) Computer, Networks, Mobility and Modeling laboratory, Department of Mathematics and Computer, FST, Hassan 1st University, Settat, Morocco*

*(2) e-NGN Research group, Africa and Middle East*

*Abstract* — WiMAX (Worldwide Interoperability for Microwave Access) technology has emerged in response to the increasing demand for multimedia services in the internet broadband networks. WiMAX standard has defined five different scheduling services to meet the QoS (Quality of Service) requirement of multimedia applications and this paper investigates one specific scheduling service, i.e. UGS scheduling. In parallel, it was observed that in the difference of the traditional quality assessment approaches, nowadays, current researches are centered on the user perception of the quality, the existing scheduling approaches take into account the QoS, mobility and many other parameters, but do not consider the Quality of Experience (QoE). In order to control the packet transmission rate so as to match with the minimum subjective rate requirements of each user and therefore reduce packet loss and delays, an efficient scheduling approach has been proposed in this paper. The solution has been implemented and evaluated in the WiMAX simulation platform developed based on NS-2. Simulation results show that by applying various levels of MOS (Mean Opinion Score) the QoE provided to the users is enhanced in term of jitter, packet loss rate, throughput and delay.

*Keywords:* WiMAX, QoE, QoS, UGS, NS-2.

## **2. Paper 31081414: Discovering Yearly Fuzzy Patterns (pp. 7-12)**

*F. A. Mazarbhuiya, College of Computer Science and IT, Albaha University, Albaha, KSA*

*Abstract* — Extracting fuzzy patterns from temporal datasets is an interesting data mining problems. An example of such pattern is yearly fuzzy pattern where a pattern holds in a certain fuzzy time interval of every year. It involves finding frequent sets and then association rules that holds in certain fuzzy time intervals (late summer or early winter etc.) in every year. In most of the previous works, the fuzziness is user-specified. However, in some applications, user may not have enough prior knowledge about the datasets under consideration and may miss some fuzziness associated with the problem. It may also the case that user may not be able to specify the fuzziness due to limitation of natural language. In this paper, we propose a method of extracting patterns that holds in certain fuzzy time intervals of every year where fuzzy time interval is generated by the method itself. The efficacy of the method is demonstrated with experimental results.

*Keywords-* Frequent itemsets, Superimposition of time intervals, Fuzzy time intervals, Right reference functions, left reference functions, Membership functions.

## **3. Paper 31081417: Assessment of Non Transmittable Codewords Enhancement to Viterbi Algorithm Decoding (pp. 13-18)**

*Salehe I. Mrutu (1), Anael Sam (2) and Nerey H. Mvungi (3)*

*(1, 2) School of Computational and Communication Science and Engineering (CoCSE), Nelson Mandela African Institution of Science and Technology (NM-AIST), Arusha, Tanzania*

*(3) College of Information and Communication Technologies, University of Dar Es Salaam (UDSM), Dar Es Salaam, Tanzania*

*Abstract* — Researchers have shown that practical mobile communication channels introduce errors that are concentrated in a certain locality rather than random errors. These are burst errors caused by deep fading of the wireless channel or a lightning strike. The existing Viterbi Algorithm (VA) capable of correcting random errors is inefficient in correcting burst errors and therefore resulting in unacceptable amount of residual errors. This paper presents an assessment of Non-Transmittable Codewords (NTCs) enhancement technique to VA in decoding the received signals subjected to burst errors that may occur in poor channels. A hard decision, 1/2 rate and constraint length K is equal to 3 Viterbi Algorithm decoding technique, Binary Phase-Shift Keying (BPSK) and Additional White Gaussian Noise (AWGN) are components used in MATLAB software based simulation when assessing the proposed technique. Applying 6NTCs to VA decoder enables the decoder to reduce 83.7 percent of its residual errors. However, the technique reduces the encoder's data transmission rate from 1/2 to 1/6.

*Keywords*-Locked Convolutional encoder; Bust errors; Residual errors; Non-Transmittable Codewords (NTCs); Viterbi Algorithm Decoding; Data Interleaver

#### **4. Paper 31081418: An Integrated Digital Academic Repository Model for Higher Learning Institutions (Case of Tanzania) (pp. 19-27)**

*Martha Mhongole, School of computational communication science and engineering, NM-AIST, Arusha, Tanzania*  
*Loserian Laizer, School of computational communication science and engineering, NM-AIST, Arusha, Tanzania*

*Abstract* — This paper explores the current existing models and technologies used in knowledge creation, knowledge sharing and knowledge dissemination practices in Higher Learning Institutions (HLIs) of Tanzania and proposes the model for the development of an Integrated Digital Academic Repository that enhances management, sharing and dissemination of Scholarly works produced in HLIs of Tanzania. The proposed model is presented and described in the paper. The study was carried out in three HLI using questionnaires, interview, observation and review of literatures. The findings show that, universities produce wide range of intellectual outputs such as research articles, learning materials, theses and technical reports. More than half population involved in the study create and store their intellectual outputs in personal computer hard drives while others store in internet cloud servers and departmental web servers. Moreover, sharing and dissemination of Intellectual output is done through internet i.e. Emails, social network, institution website and cloud servers, journal publication, seminar presentations, posters and printed copies in libraries. The identified methods proven to be unreliable and hindering availability and accessibility of scholarly works. Thus the proposed model provides a central system through which intellectual outputs will be collected, organized and archived and disseminated through it. The paper concludes with the conceptual framework of the proposed system, whereas design and development carried forward to be our future work.

*Keywords*- Higher learning institution, intellectual output, knowledge management, knowledge sharing, model, digital repository

#### **5. Paper 31081425: Finding Untraced Fuzzy Association Rules (pp. 28-30)**

*F. A. Mazarbhuiya, College of Computer Science and IT, Albaha University, KSA*

*Abstract* — Fuzzy association rules are rules of the form “If X is A then Y is B” where X and Y are set of attributes and A, B are fuzzy sets that describe X and Y respectively. In most of fuzzy association rules mining problem fuzziness is specified by users. The users usually specify the fuzziness based on their understanding of the problem as well as the ability to express the fuzziness by natural language. However there exist some fuzziness which cannot be expressed using natural language due its limitation. In this paper we propose a method of extracting fuzzy association rules which cannot be traced by usual methods. We suggest a way of extracting these rules.

*Keywords*- Fuzzy set, Association rules, Fuzzy interval, Certainty factor, Significance factor, Between Operation.



## **6. Paper 31081428: A Novel Approach to Address Information Leakage Attacks Based on Machine Virtualization (pp. 31-42)**

*Omar Hussein, Nermin Hamza, Hesham Hefny*

*Computer and Information Sciences Department, Institute of Statistical Studies and Research, Cairo University, Egypt*

**Abstract** — In a traditional non-virtualized computer system the whole software stack is highly vulnerable to security breaches. This is mainly caused by the coexistence of deployed security systems in the same space as the potentially compromised operating system and applications that often run with administrative privileges. In such a structure, compromising, bypassing, disabling, or even subverting deployed security systems become trivial. Machine virtualization provides a powerful abstraction for addressing information security issues. Its isolation, encapsulation, and partitioning properties can be leveraged to reduce computer systems' susceptibility to security breaches. This paper demonstrates that machine virtualization when employed and synthesized with cryptography would preserve information confidentiality even in an untrusted machine. It presents a novel information security approach called Virtualized Anti-Information Leakage (VAIL). Its objective is to thwart malicious software and insiders' information leakage attacks on sensitive files after decryption in potentially compromised computer systems. VAIL's defenses are evaluated against a variety of information leakage attacks including: (1) direct attacks launched on sensitive files from an untrusted virtual machine, and a compromised virtual machine monitor; and (2) indirect attacks exploiting covert storage and timing channels. Based on the security evaluation, it is concluded that VAIL effectively complied with the security requirements, and met its objective.

**Index Terms**—*Information Security; Information Leakage; Machine Virtualization; Malicious Software; Insider Threat*

## **7. Paper 31081429: Performance Analysis of Speech Quality in VoIP during Handover (pp. 43-48)**

*M. Yousef & M. Fouad*

*Electronics & Communications Dept., Faculty of Eng., Zagazig Uni., Egypt.*

**Abstract** — Quality of Service is a very important factor to determine the quality of a VoIP call. Different subjective and objective models exist for evaluating the speech quality in VoIP. E-model is one of the objective methods of measuring the speech quality; it considers various factors like packet loss, delay and codec impairments. The calculations of E-model are not very accurate in case of handovers – when a VoIP call moves from one wireless LAN to another. This paper conducted experimental evaluation of performance of E-model during handovers and proposes a new approach to accurately calculate the speech quality of VoIP during handovers and make MOS calculator which take the results through. A detailed description of the experimental setup and the comparison of the new approach with E-model is presented in this work.

## **8. Paper 31081433: Web Users Clustering Analysis (pp. 49-52)**

*Hooman Rokham, Computer Engineering Department, Shahre-e-Qods Branch, Islamic Azad University, Shahr-e-Qods, Iran*

*Hale Falakshahi, Computer Engineering Department, Science and Research Branch, Islamic Azad University, Neyshabur, Iran*

**Abstract** — As one of the most important tasks of web usage mining, web user clustering, which establishes groups of users exhibiting similar browsing patterns, provides useful knowledge to personalized web services. There are many clustering algorithms. In this paper, users' similarity is calculated then a comparative analysis of two clustering algorithms namely K-means algorithm and hierarchical algorithm is performed. Web users are clustered with these algorithms based on web user log data. Given a set of web users and their associated historical web usage data, we study their behavior characteristic and cluster them. In terms of accuracy K-means produces better results as compared to hierarchical algorithm.



*Keywords-clustering; K-means algorithm; hierarchical algorithm*

## **9. Paper 31081440: A Comparative Analysis of Dynamic Scheduling Algorithms versus the Round-Robin Scheduling Algorithm (pp. 53-60)**

*Vilma Tomço, University of Tirana, Faculty of Mathematics, Statistics and Applied Informatics, Tirana, Albania*

*Anduela Dervishi, Polytechnic University of Tirana, Faculty of Mathematical and Physical Engineering, Tirana, Albania*

*Elton Lika, Polytechnic University of Tirana, Faculty of Information Technology, Department of Informatics Engineering, "Mother Teresa" Square, Tirana, Albania*

*Igli Tafa, Polytechnic University of Tirana, Faculty of Information Technology, Department of Informatics Engineering*

*Abstract* - Scheduling is one of the most important concepts in Operating Systems. One of the most popular algorithms is Round - Robin, which switches the processes after running the set Time Quantum (TQ). TQ value affects the average time of Waiting and Turnaround, and the number of Context Switches (CS). This definition can be static, which does not change, and dynamic, calculated cycle after cycle. This review builds on the study of new techniques for the determination of TQ dynamically. Initially is shown that in all cases this method is efficient and then we rank the most important techniques used. We look at how each works and the differences and their similarities. We will observe their efficiency in different parameters and the conditions in which they are effective. Finally we show that MDTQRR is most effective, minimizing the number of CS and Harm is the most effective in AVG (Waiting and Turnaround) Time.

*Key words* - Round-Robin, Quantum Time, Waiting Time, Turnaround Time, Context Switch, ready queue.

# An Improved UGS Scheduling with QoE Metrics in WiMAX Network

Tarik ANOUARI<sup>1</sup>

Abdelkrim HAQIQ<sup>1,2</sup>

1 Computer, Networks, Mobility and Modeling laboratory  
Department of Mathematics and Computer  
FST, Hassan 1st University, Settat, Morocco  
2 e-NGN Research group, Africa and Middle East

**Abstract**— WiMAX (Worldwide Interoperability for Microwave Access) technology has emerged in response to the increasing demand for multimedia services in the internet broadband networks. WiMAX standard has defined five different scheduling services to meet the QoS (Quality of Service) requirement of multimedia applications and this paper investigates one specific scheduling service, i.e. UGS scheduling. In parallel, it was observed that in the difference of the traditional quality assessment approaches, nowadays, current researches are centered on the user perception of the quality, the existing scheduling approaches take into account the QoS, mobility and many other parameters, but do not consider the Quality of Experience (QoE). In order to control the packet transmission rate so as to match with the minimum subjective rate requirements of each user and therefore reduce packet loss and delays, an efficient scheduling approach has been proposed in this paper. The solution has been implemented and evaluated in the WiMAX simulation platform developed based on NS-2. Simulation results show that by applying various levels of MOS (Mean Opinion Score) the QoE provided to the users is enhanced in term of jitter, packet loss rate, throughput and delay.

**Keywords:** WiMAX, QoE, QoS, UGS, NS-2.

## I. INTRODUCTION

Habitually, the network has been assessed objectively by measuring some parameters to evaluate the network service quality. This evaluation is known as the QoS of the network. The term QoS refers to the guarantees on the ability of a network to deliver predictable results and a more deterministic performance, so data can be transferred with a minimum delay, packet loss, jitter and maximum throughput. The QoS does not take into account the user's perception of the quality. Another approach which takes into account the user's perception is named QoE, it's the overall acceptability of an application or service, as perceived subjectively by the end user, it groups together user perception, expectations, and experience of application and network performance.

In order to get a more comprehensive view of the quality perceived by end users, QoE it has become increasingly a very interesting area of research. Many related works was presented on analyzing and improving QoE [12] in WiMAX network. The study presented in [14] suggested an estimation method of QoE metrics based on QoS metrics in WiMAX network. The QoE was estimated by using a Multilayer Artificial Neural Network (ANN). The results show an efficient estimation of metrics of QoE with respect to QoS parameters.

In [6, 7, 8], the authors focus on the ANN method to adjust the input network parameters to get the optimum output to satisfy the end users. Especially, the success of the ANN approach depends on the model's capacity to completely learn the nonlinear interactions between QoE and QoS. In [16], Muntean proposes a learner QoE model that considers delivery performance-based content personalization in order to improve user experience when interacting with an online learning system. Simulation results show significant improvements in terms of learning achievement, learning performance, learner navigation and user QoE.

In [3], our study was focused on studying and analyzing QoS performances of VoIP traffic using different service classes in term of throughput, jitter and delay. The simulation results show that UGS service class is the best suited to handle VoIP traffic. This paper proposes a QoE-based model in order to provide best performances in WiMAX network especially for the real-time traffic. The target of this improvement is to schedule traffic of UGS service class.

The rest of this paper is organized as follows. A short description of WiMAX technology is given in section 2. In section 3, a QoE overview background is presented. The proposed QoE model is described in detail in section 4. Simulation environment and performance parameters are presented in section 5. Section 6 shows simulation results and analysis. Finally, section 7 concludes the paper.

## II. WiMAX TECHNOLOGY

WiMAX is a wireless communication standard based on the 802.16 standards [10, 11], the main objective of WiMAX is to provide an Internet broadband connection to a coverage area with a radius of several kilometers. Unlike ADSL (Asymmetric Digital Subscriber Line) or other wired technologies, WiMAX uses radio waves, similar to those used for mobile phone.

WiMAX can be used in point-to-multipoint (PMP) mode in which serving multiple client terminals is ensured from a central base station, and in point-to-point (P2P) mode, in which there is a direct link between the central base station and the subscriber.

PMP mode is less expensive to implement and operate while P2P mode can provide greater bandwidth.

### A. QoS in WiMAX Network

Since QoS support is an important part of WiMAX network, the concept of QoS was introduced natively in WiMAX [18], so this protocol ensures the good operation of a service. Some services are very demanding; VoIP cannot tolerate delay in the transmission of data. WiMAX uses service classes to allow different QoS between each communication.

The concept of QoS mainly depends on the service provided, its sensitivity to transmission errors, its requirement of response time... etc. For VoIP traffic, one of the challenges is related to network congestion and latency, we will need a real-time traffic transfer, with very low latency and low jitter. A complete definition of QoS often refers to the mode of transport of information, although the solution adopted by the network to provide the service must remain transparent to the user.

Satisfying QoS requirement becomes very imperative in IEEE802.16 systems to provide best performance, in particular in the presence of various types of connections, namely the current calls, new calls and the handoff connection.

### B. WiMAX Network Architecture

The architecture of WiMAX network consists of base station named BTS (Base Transceiver Station) or BS (Base Station) and mobile clients or stations (SS Subscriber Station). The base station acts as a central antenna responsible for communicating and serve mobile stations, in their turn, serve clients using WIFI or ADSL. The BS can provide various levels of QoS over its queuing, scheduling, control, signaling mechanisms, classification and routing. Figure 1 shows the architecture of WiMAX network [10, 11].

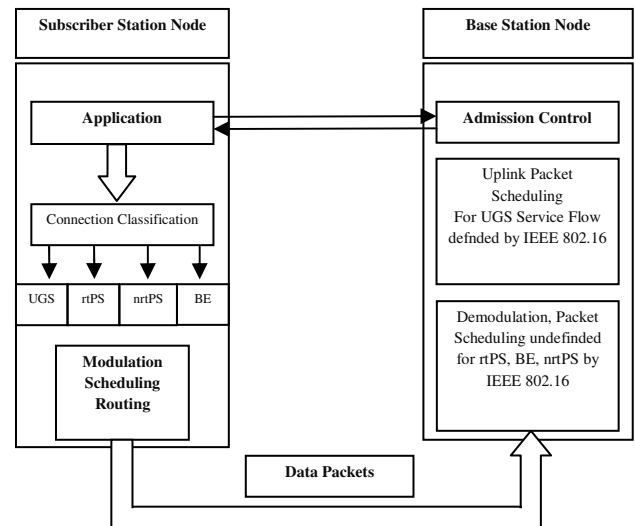


Figure 1: WiMAX Network Architecture

### C. Different Service Classes in WiMAX

Multiple kinds of traffic are considered in WiMAX. QoS is negotiated at the service flow, especially at the establishment of the connection. A modulation and coding technique are set up. To satisfy different types of applications, WiMAX standard has defined four service classes of quality, namely Unsolicited Grant Service (UGS), Best Effort (BE), real-time Polling Service (rtPS) and non-real time Polling Service (nrtPS). The amendment to the IEEE 802.16e standard (802.16e 2005) [1] on mobility includes a fifth type of service class, the extended real-time Polling Service (ertPS). This service is placed between the UGS service and rtPS service. It can serve real-time applications that generate periodic packets of variable size, the example given in the standard is that of a VoIP application with silence suppression.

Some services like VoIP are very demanding in term of QoS, it cannot tolerate delay in data transmission while others have fewer requirements.

Table 1 classifies different service classes of WiMAX and gives their description and QoS parameters.

TABLE I. SERVICE CLASSES IN WiMAX

Service	Description	QoS parameters
UGS	Real-time data streams comprising fixed size data packets at periodic intervals	Maximum Sustained Rate Maximum Latency Tolerance Jitter Tolerance
rtPS	support real-time service flows that periodically generate variable-size data packets	Traffic priority Maximum latency tolerance Maximum reserved rate

ertPS	Real-time service flows that generate variable-sized data packets on a periodic basis.	Minimum Reserved Rate Maximum Sustained Rate Maximum Latency Tolerance Jitter Tolerance Traffic Priority
nrtPS	Support for non-real-time services that require variable size data grants on a regular basis	Traffic priority Maximum reserved rate Maximum sustained rate
BE	Data streams for which no data minimum service level is required.	Maximum Sustained Rate Traffic Priority

### III. QUALITY OF EXPERIENCE

Quality of Experience (QoE, user Quality of Experience or simply QX) is a subjective measure that reflects the user satisfaction with the service provided (web browsing, phone call, TV broadcast, call to a Call Center).

Today, assessing the quality of experience has become essential for service providers and content providers.

#### A. Quality of Experience vs Quality of Service assessment

QoS appeared in the 90 years to describe the quality of the network. Since that time the acronym QoS has been usually used to describe the improved performance realized by hardware and / or software. But with the rapid improvement of Media services, this measure has shown its limitations and many efforts have been made to develop a new metric that reflects more accurately the quality of service provided. This measure is called the QoE.

QoE is a subjective measure of a customer's experiences with a service according to his perception. Indeed, the notion of user experience has been introduced for the first time by Dr. Donald Norman, citing the importance of designing a user centered service [17].

Gulliver and Ghinea [9] classify QoE into three components: assimilation, judgment and satisfaction. The assimilation is a quality assessment of the clarity of the contents by an informative point of view. The judgment of quality reflects the quality of presentation. Satisfaction indicates the degree of overall assessment of the user.

QoE and QoS have become complementary concepts: QoS indicators are used to identify and analyze the causes of network congestions while QoE indicators are used to monitor the quality offered to users. These two solutions used in parallel are a complete system monitoring.

#### B. QoE measurement approaches

Two main quality evaluation methodologies are defined, namely objective and subjective performance evaluation. Subjective assessments are carried out by

end users who are asked to evaluate the overall perceived quality of the service provided, the most frequently used measurement is the MOS recommended by the International Telecommunication Union (ITU) [13], and it's defined as a numeric value evaluation from 1 to 5 (i.e. poor to excellent).

Objective methods are centered on algorithms, mathematical and/or comparative techniques that generate a quantitative measure of the service provided.

Peter and Bjørn [5] classified the existing approaches of measuring network service quality from a user perception into three classifications, namely: Testing User-perceived QoS (TUQ), Surveying Subjective QoE (SSQ) and Modeling Media Quality (MMQ). The first two approaches collect subjective information from users, whereas the third approach is based on objective technical assessment. Figure 2 [2] gives an overview of the classification of the existing approaches.

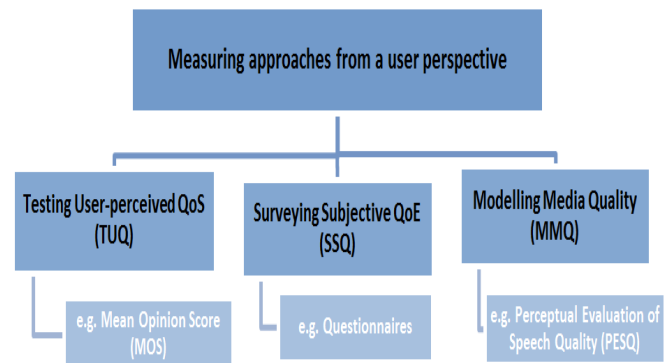


Figure 2. The approaches for measuring network service quality from a user perception

### IV. QoE-BASED SCHEDULING ALGORITHM MODEL

In this section, we propose a QoE-based scheduling approach in WiMAX network, because it's observed that the existing scheduling algorithms take into account QoS but not user perception of the service provided, where every user has different subjective requirement of the system.

#### A. Proposed QoE model

In the proposed QoE-based model three QoE levels are used, each user has an initial maximum transmission rate, a minimum subjective rate requirement and a subjective threshold value. The traffic starts with a maximum transmission rate on each user. When the packet loss rate is greater than the user selected threshold (which is chosen at the beginning of the simulation), then each user checks if the transmission rate is higher than the minimum subjective requirement, if yes the transmission rate is decreased, otherwise it's remained at the same level.

In the other hand, if the packet loss rate is less than the selected threshold, then the user checks if the transmission rate is lower than the minimum subjective requirement, if yes the transmission rate is increased, otherwise it's remained at the same level.

The threshold can be selected by the user as a percentage of the data transmission rate, for example, if the user introduces a value of 50 as a threshold then the threshold for packet loss rate is 50%. Figure 3 shows the activity diagram of the proposed model.

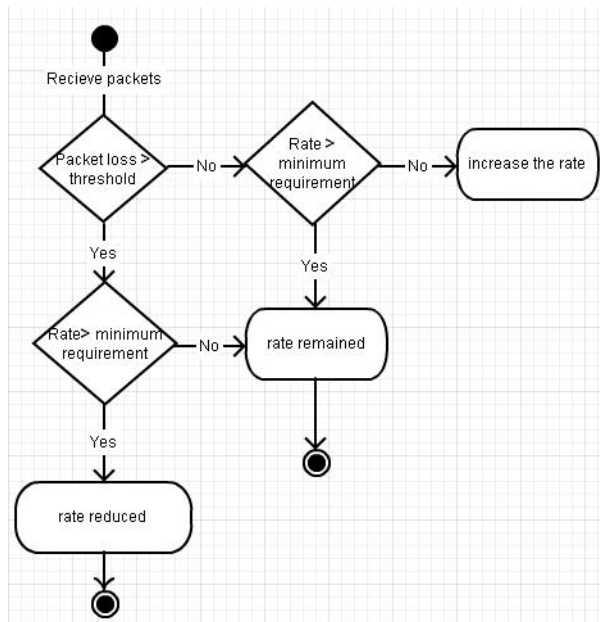


Figure 3: Activity diagram of the proposed QoE-Model

## V. SIMULATION ENVIRONNEMENT

### A. Simulation Model

In this paper, we evaluate the performances of the proposed QoE-based scheduling algorithm, as we consider the Wireless-OFDM PHY layer, our QoE-model is evaluated and compared with the popular WiMAX module developed by NIST (National Institute for Standards and Technologies), which is based on the IEEE 802.16 standard (802.16-2004) and the mobility extension (802.16e-2005) [19]. Our simulation scenario consists of creating five wireless users connected to a base station (BS). A sink node is created and attached to the base station to accept packets. A traffic agent is created and then attached to the source node. The Network Simulator (NS-2) [15] is used.

Finally, we set the traffic that produces each node. The first node has run with CBR (Constant Bit Rate) packet size of 200 bytes and interval of "0,0015", the second node has run with CBR packet size of 200 bytes and interval of "0,001", the third node has run with CBR packet size of 200 bytes and interval of "0,001", the fourth node has run with CBR packet size of 200 bytes

and interval of "0,001" and fifth node has run with CBR packet size of 200 bytes and interval of "0,0015". The initial transmission rate that produces each node is about "133,3 Kbps", "200 Kbps", "200 Kbps", "200 Kbps" and "133,3 Kbps" respectively. All nodes have the same priority.

Each user has a minimum requirement, so the first user requires minimal traffic rate of "120 Kbps", the second "150 Kbps", the third "150 Kbps", the fourth "150 Kbps" and the fifth "120 Kbps".

The following table summarizes the above description about the produced and required traffic rate of each user.

TABLE II. USER'S TRAFFIC PARAMETERS

Traffic rate Users	Initial traffic rate (Kbps)	User minimum requirement (Kbps)
User 1	133,33 (200byte/0. 0015)	120
User 2	200 (200byte/0. 001)	150
User 3	200 (200byte/0. 001)	150
User 4	200 (200byte/0. 001)	150
User 5	133.33 (200byte/0. 0015)	120

We use five different thresholds 10%, 20%, 30%, 40% and 50%.

We have used the QoS-included WiMAX module [4] within NS-2.29. This module is based on the NIST implementation of WiMAX [19], it includes the QoS classes as well as the management of the QoS requirements, unicast and contention request opportunities mechanisms, and scheduling algorithms for the UGS, rtPS and BE QoS classes.

The resulted trace files are interpreted and analyzed based on a PERL script, which is an interpretation script software used to extract data from trace files to get throughput, packet loss rate, jitter and delay. The extracted results are plotted in graphs using EXCEL software.

### B. Simulation Parameters

The same simulation parameters are used for both NIST and QOE-based scheduling algorithms, table 3 summarizes the simulation parameters:

TABLE III. SIMULATION PARAMETERS

Parameter	Value
Network interface type	Phy/WirelessPhy/OFDM
Propagation model	Propagation/OFDM
MAC type	Mac/802.16/BS
Antenna model	Antenna/OmniAntenna
Service class	UGS
packet size	200 bytes
Frequency bandwidth	5 MHz
Receive Power Threshold	2,025e-12
Carrier Sense Power Threshold	0,9 * Receive Power Threshold

Channel	3,486e+9
Simulation time	200s

### C. Performance Parameters

Main QoS parameters were analyzed in our simulation, namely average throughput, packet loss rate, average delay and average jitter.

## VI. SIMULATION RESULTS AND ANALYSIS

We have performed various simulation scenarios in order to analyse and compare the proposed QoE-based scheduler with the NIST scheduler in term of average throughput, packet loss rate, average delay and average jitter in WiMAX network using UGS service class.

In figure 5, we note that the average throughput in the case of the QoE-based scheduler algorithm is lower than for the NIST scheduler for all flows, whereas the third flow has the largest range between maximum and minimum values.

For the flows 2 and 4 the throughput values are similar for both NIST scheduler and QoE-based scheduler, especially when the QoE threshold is 50%.

The scheduler that takes into account the QoE varied the throughput for different users so as to match with the minimum subjective rate requirements of each user in order to reduce jitter, delays and packet loss.

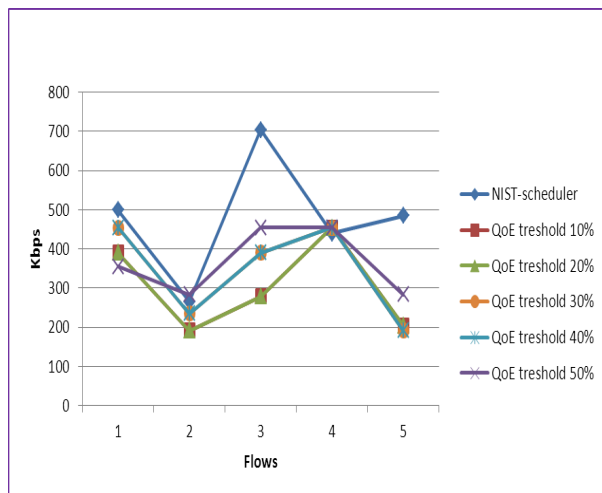


Figure 5. Average throughput

The improvement is noticeable as shown in Figure 6 when the QoE-based scheduler is used. The packet loss rate for all users is reduced while the packet loss rate is similar for both schedulers in the case of flows 3 and 5. The NIST scheduler gives lower performances compared with the QoE-based scheduler in term of packet loss rate.

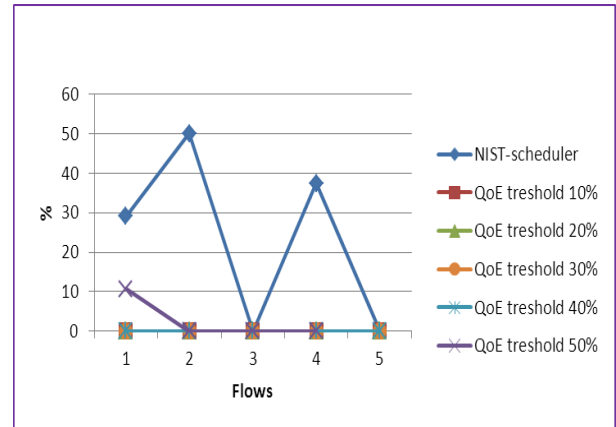


Figure 6. Packet loss rate

It can be observed from the figure 7 that the proposed QoE-based scheduler algorithm has lowest values of average jitter compared with the NIST scheduler by applying different threshold levels, especially for the flows 1, 2 and 3. Average jitter values are identical for flows 4 and 5 for all the threshold levels.

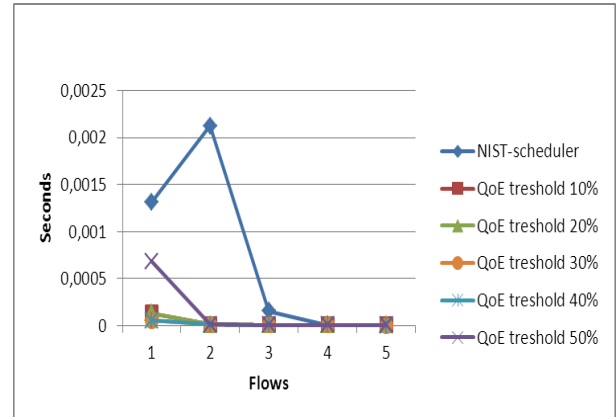


Figure 7. Average Jitter

As shown in figure 8, the QoE-based scheduler outperforms the NIST scheduler, the average transmission packets delay values still lowest in the case of QoE-scheduler, while the two schedulers have similar values for flows 4 and 5.

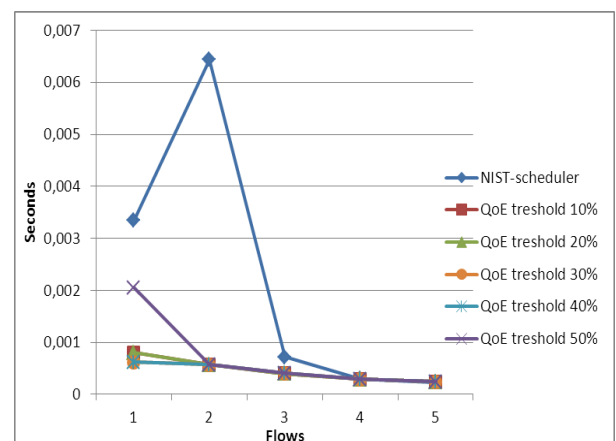


Figure 8. Average Delay

## VII. CONCLUSION

In this paper, we have proposed a new QoE-based scheduler in order to manage the packet transmission rate for users in WiMAX network. When the packet loss rate exceeds some threshold, there are two cases, either the transmission packet rate is less than the minimum subjective rate requirement, then the user continue to transmit with the same packet transmission rate, otherwise he should reduce it.

The simulations carried out show that the use of different levels of MOS improves the QoE provided to users of WiMAX network. The proposed QoE-model significantly reduced packet loss, delay and jitter, the transmission rate is reduced for each connection, until matching with its minimum subjective rate requirement.

As a future work we may extend this study by adding other parameters like mobility models.

## REFERENCES

- [1] Z. Abichar, Y. Peng and J. Morris Chang, "WiMax: The Emergence of Wireless Broadband", IT Professional, Vol. 8, Issue. 4, pp. 44-48, Doi:10.1109/MITP.2006.99, July-Aug. 2006.
- [2] M. Alreshoodi, J. Woods, "Survey on QoE/QoS Correlation Models for Multimedia Services", International Journal of Distributed and Parallel Systems (IJDPs) Vol.4, No.3, May 2013.
- [3] T. Anouari and A. Haqiq, "Analysis of VoIP and Video Traffic over WiMAX Using Different Service Classes", Journal of Mobile Multimedia, Vol. 9, No.3&4, pp. 230-241, 2014.
- [4] A. Belghith, L. Nuaymi "Design and Implementation of a QoS-included WiMAX Module for NS-2 Simulator", SIMUTools 2008, Marseille, France, March 3-7, 2008.
- [5] P. Brooks, B. Hestnes, "User measures of quality of experience: Why being objective and quantitative is important". IEEE Network 24(2): pp. 8-13, 2010.
- [6] P. Calyam, P. Chandrasekaran, G. Trueb, N. Howes, R. Ramnath, D. Yu, Y. Liu, L. Xiong, & D. Yang, "Multi-Resolution Multimedia QoE Models for IPTV Applications", Volume 2012, Article ID 904072, 13 pages doi:10.1155/904072, 2012.
- [7] H. Du, C. Guo, Y. Liu & Y. Liu, (2009) "Research on Relationships between QoE and QoS based on BP Neural Network", In: Proceedings of IC-NIDC, pp. 312-315, 2009.
- [8] P. Frank & J. Incera, "A neural network based test bed for evaluating the quality of video streams in IP networks", 0-7695-2569-5/06 © IEEE, Proceedings of the Electronics, Robotics and Automotive Mechanics Conference (CERMA'06), 2006.
- [9] S. R. Gulliver and G. Ghinea. "The Perceptual and Attentive Impact of Delay and Jitter in Multimedia Delivery". IEEE Transactions on Broadcasting, 53(2): pp. 449-458, June 2007.
- [10] IEEE 802.16-2004, "IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems", October, 2004.
- [11] IEEE standard 802.16-2005, "IEEE standard for Local and Metropolitan Area Networks-Part16: Air Interface for Fixed and Mobile Broadband wireless Access systems Amendment 2", February 28, 2006.
- [12] ITU-T Rec. P.10/G.100, Amendment 2, "New definitions for inclusion in Recommendation ITU-T P.10/G.100," July 2008.
- [13] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality". <http://www.itu.int/>, Geneva, 08/1996
- [14] V. Machado, C. Oliveira, A. Marcelino, S. Carlos, N. Vijaykumar, C. Hirata, "A New Proposal to Provide Estimation of QoS and QoE over WiMAX Networks, An approach based on computational intelligence and discrete-event simulation", 978-1-4673-0279-1©IEEE, 2011.
- [15] Marc Greis, "Tutorial for Network Simulator NS", <http://www.scribd.com/doc/13072517/tutorial-NS-full-byMARC-GREIS>.
- [16] C. H. Muntean, "Improving learner quality of experience by content adaptation based on network conditions", Computers in Human Behavior, 24(4), pp. 1452-1472, 2007.
- [17] D. Norman and S. Draper, "User centered system design: New perspectives on human-computer interaction". L. Erlbaum Associates, 1986.
- [18] P. Rengaraju, C.H. Lung, A. Srinivasan, R.H.M. Hafez, "Qos Improvements in Mobile WiMAX Networks", AHU J. of Engineering & Applied Sciences, Vol. 3, Issue 1, pp. 107-118, 2010.
- [19] Seamless and secure mobility. [http://www.nist.gov/itl/antd/emntg/ssm\\_tools.cfm](http://www.nist.gov/itl/antd/emntg/ssm_tools.cfm)



# Discovering Yearly Fuzzy Patterns

F. A. Mazarbhuiya  
College of Computer Science and IT  
Albaha University  
Albaha, KSA

**Abstract**— Extracting fuzzy patterns from temporal datasets is an interesting data mining problems. An example of such pattern is yearly fuzzy pattern where a pattern holds in a certain fuzzy time interval of every year. It involves finding frequent sets and then association rules that holds in certain fuzzy time intervals (late summer or early winter etc.) in every year. In most of the previous works, the fuzziness is user-specified. However, in some applications, user may not have enough prior knowledge about the datasets under consideration and may miss some fuzziness associated with the problem. It may also the case that user may not be able to specify the fuzziness due to limitation of natural language. In this paper, we propose a method of extracting patterns that holds in certain fuzzy time intervals of every year where fuzzy time interval is generated by the method itself. The efficacy of the method is demonstrated with experimental results.

**Keywords**- *Frequent itemsets, Superimposition of time intervals, Fuzzy time intervals, Right reference functions, left reference functions, Membership functions.*

## I. INTRODUCTION

Among the various types of data mining applications, analysis of transactional data has been considered important. It is assumed that the dataset keeps information about users transactions. In a market-basket data set each transaction is a collection of items bought by a customer at one time. The notion proposed in [1] is to capture the co-occurrence of items in transactions, given two percentage parameters as minimum support and minimum confidence thresholds. One important extension above-mentioned problem is to include a time attribute. When a customer buy something, this transaction and it's time of transaction is automatically recorded. In [2], *Ale et al* has proposed a method of discovering association rules that hold within the life-span of the corresponding item set and not within the life-span of the whole dataset.

In [3] the concept of locally frequent item sets has been proposed which are itemsets that are frequent in certain time intervals and may or may not be frequent through out the life-span of the item set. In [3] an algorithm has been proposed for finding such itemsets along with a list of sequences of time intervals. Here each frequent itemset is associated with a sequence of time intervals where it is frequent. Considering the time-stamp as calendar dates a method is discussed in [4] which can extract yearly, monthly and daily periodic or partially periodic patterns. If the periods are kept in a compact manner using the method discussed in [4], it turns out to be a fuzzy time interval. In [4], the author put a restriction that the

intervals to be superimposed must have overlapping upto a certain specified extent. In this paper, we discuss such patterns and devise algorithms for extracting such patterns. The algorithm can be applied for extracting monthly or daily fuzzy patterns also. Although our algorithm is quite similar to the algorithm discussed in [4], but in our case we removed the restriction on the intervals to be superimposed. We superimposed all the intervals that have non-empty intersection which turns out to be fuzzy intervals. The paper is organized as follows. In section-II, we discuss related works. In section-III, we discuss terms, definitions and notations used in the algorithm. In section-IV, the proposed algorithm is discussed along with complexity. In section-V, we discuss about results and analysis. Finally a summary and lines for future works are discussed in section-VI.

## II. RELATED WORKS

The problem of discovery of association rules was first formulated by *Agrawal et al* [1]. Given a set  $I$ , of items and a large collection  $D$  of transactions involving the items, the problem is to find relationships among the presence of various items in the transactions..

Temporal Data Mining [5] is an important extension of conventional data mining. By taking into account the time aspect, more interesting patterns that are time dependent can be extracted. The association rule discovery process is also extended to incorporate temporal aspects. The problems associated are to find valid time periods during which association rules hold and the discovery of possible periodicities that association rules have. In [2] an algorithm for the discovery of temporal association rules is described. For each item (which is extended to item set) a lifetime or life-span is defined which is the time gap between the first occurrence and the last occurrence of the item in the transaction in the database. Supports of items are calculated only during its life-span. Thus each rule has associated with it a time frame. In [3], the works done in [2] has been extended by considering time gap between two consecutive transactions containing an item set into account.

Considering the periodic nature of patterns, *Ozden* [6] proposed a method, which is able to find patterns having periodic nature where the period has to be specified by the user. In [7], *Li et al* the authors discuss about a method of extracting temporal association rules with respect to *fuzzy match* i.e. association rule holding during “enough” number of intervals given by the corresponding calendar pattern. Similar works were done in [8] incorporating multiple granularities of time

intervals (e.g. first working day of every month) from which both cyclic and user defined calendar patterns can be achieved.

Mining fuzzy patterns from datasets have been studied by different authors. In [9], the authors present an algorithm for mining fuzzy temporal patterns from a given process instance. Similar work is done in [10]. In [11] method of extracting fuzzy periodic association rules is discussed.

### III. TERMS DEFINITIONS AND NOTATIONS USED

Let us review some definitions and notations used in this paper.

Let  $E$  be the universe of discourse. A fuzzy set  $A$  in  $E$  is characterized by a membership function  $A(x)$  lying in  $[0, 1]$ .  $A(x)$  for  $x \in E$  represents the grade of membership of  $x$  in  $A$ . Thus a fuzzy set  $A$  is defined as

$$A = \{ (x, A(x)), x \in E \}$$

A fuzzy set  $A$  is said to be normal if  $A(x) = 1$  for at least one  $x \in E$ .

A fuzzy number is a convex normalized fuzzy set  $A$  defined on the real line  $R$  such that

1. there exists an  $x_0 \in R$  such that  $A(x_0) = 1$ , and
2.  $A(x)$  is piecewise continuous.

Thus a fuzzy number can be thought of as containing the real numbers within some interval to varying degrees.

Fuzzy intervals are special fuzzy numbers satisfying the following.

1. there exists an interval  $[a, b] \subset R$  such that  $A(x_0) = 1$  for all  $x_0 \in [a, b]$ , and
2.  $A(x)$  is piecewise continuous.

A fuzzy intervals can be thought of as a fuzzy number with a flat region. A fuzzy interval  $A$  is denoted by  $A = [a, b, c, d]$  with  $a < b < c < d$  where  $A(a) = A(d) = 0$  and  $A(x) = 1$  for all  $x \in [b, c]$ .  $A(x)$  for all  $x \in [a, b]$  is known as *left reference function* and  $A(x)$  for  $x \in [c, d]$  is known as the *right reference function*. The *left reference function* is non-decreasing and the *right reference function* is non-increasing.

The *support* of a fuzzy set  $A$  within a universal set  $E$  is the crisp set that contains all the elements of  $E$  that have non-zero membership grades in  $A$  and is denoted by  $S(A)$ . Thus

$$S(A) = \{ x \in E; A(x) > 0 \}$$

The *core* of a fuzzy set  $A$  within a universal set  $E$  is the crisp set that contains all the elements of  $E$  having membership grades 1 in  $A$ .

#### Set Superimposition

When we overwrite, the overwritten portion looks darker for obvious reason. The set operation union does not explain this phenomenon. After all

$$A \cup B = (A - B) \cup (A \cap B) \cup (B - A)$$

and in  $(A \cap B)$  the elements are represented once only.

In [13] an operation called *superimposition* denoted by  $(S)$  was proposed. If  $A$  is superimposed over  $B$  or  $B$  is superimposed over  $A$ , we have

$$A(S)B = (A - B) (+) (A \cap B)^{(2)} (+) (B - A) \quad \dots (1)$$

Where  $(A \cap B)^{(2)}$  are the elements of  $(A \cap B)$  represented twice, and  $(+)$  represents union of disjoint sets.

To explain this, an example has been taken.

If  $A = [a_1, b_1]$  and  $B = [a_2, b_2]$  are two real intervals such that  $A \cap B \neq \emptyset$ , we would get a superimposed portion. It can be seen from (1)

$$[a_1, b_1] (S) [a_2, b_2] = [a_{(1)}, a_{(2)}] (+) [a_{(2)}, b_{(1)}]^{(2)} (+) (b_{(1)}, b_{(2)}) \quad \dots (2)$$

where

$$a_{(1)} = \min(a_1, a_2) \quad a_{(2)} = \max(a_1, a_2)$$

$$b_{(1)} = \min(b_1, b_2), \text{ and } b_{(2)} = \max(b_1, b_2)$$

(2) explains why if two line segments are *superimposed*, the common portion looks doubly dark [5]. The identity (2) is called *fundamental identity of superimposition of intervals*.

Let now,  $[a_1, b_1]^{(1/2)}$  and  $[a_2, b_2]^{(1/2)}$  be two fuzzy sets with constant membership value  $1/2$  everywhere (i.e. equi-fuzzy intervals with membership value  $1/2$ ). If  $[a_1, b_1] \cap [a_2, b_2] \neq \emptyset$  then applying (2) on the two equi-fuzzy intervals we can write

$$[a_1, b_1]^{(1/2)} (S) [a_2, b_2]^{(1/2)} = [a_{(1)}, a_{(2)}]^{(1/2)} (+) [a_{(2)}, b_{(1)}]^{(1)} (+) (b_{(1)}, b_{(2)})^{(1/2)} \quad \dots (3)$$

To explain this we take the fuzzy intervals  $[1, 5]^{(1/2)}$  and  $[3, 7]^{(1/2)}$  with constant membership value  $(1/2)$  given in figure-1.1 and figure-1.2. Here  $[1, 5] \cap [3, 7] = [3, 5] \neq \emptyset$ .



Fig-1.1

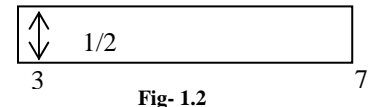


Fig- 1.2

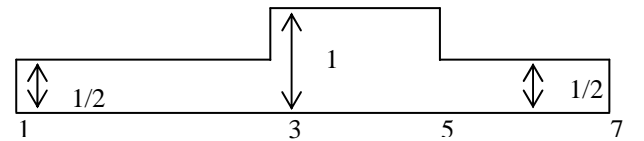


Fig-1.3 Superimposed interval

If we apply *superimposition* on the intervals then the *superimposed* interval will be consisting of  $[1, 3]^{(1/2)}$ ,  $[3, 5]^{(1)}$  and  $(5, 7]^{(1/2)}$ . Here the membership of  $[3, 5]$  is (1) due to double representation and it is shown in figure-1.3

Let  $[x_i, y_i]$ ,  $i=1,2,\dots,n$ , be  $n$  real intervals such that  $\bigcap_{i=1}^n [x_i, y_i] \neq \emptyset$ . Generalizing (3) we get

$$\begin{aligned} & [x_1, y_1]^{(1/n)} (S) [x_2, y_2]^{(1/n)} (S) \dots (S) [x_n, y_n]^{(1/n)} \\ & = [x_{(1)}, x_{(2)}]^{(1/n)} (+) [x_{(2)}, x_{(3)}]^{(2/n)} (+) \dots (+) [x_{(r)}, x_{(r+1)}]^{(r/n)} \\ & (+) \dots (+) [x_{(n)}, y_{(1)}]^{(1)} (+) [y_{(1)}, y_{(2)}]^{((n-1)/n)} (+) \dots (+) [y_{(n-r)}, y_{(n-r+1)}]^{(r/n)} \\ & (+) \dots (+) [y_{(n-2)}, y_{(n-1)}]^{(2/n)} (+) [y_{(n-1)}, y_{(n)}]^{(1/n)} \dots \quad (4) \end{aligned}$$

In (4), the sequence  $\{x_{(i)}\}$  is formed by sorting the sequence  $\{x_i\}$  in ascending order of magnitude for  $i=1,2,\dots,n$  and similarly  $\{y_{(i)}\}$  is formed by sorting the sequence  $\{y_i\}$  in ascending order.

Although the set superimposition is operated on the closed intervals, it can be extended to operate on the open and the half-open intervals in the trivial way.

**Lemma 1. (The Glivenko-Cantelli Lemma of Order Statistics)**

Let  $X = (X_1, X_2, \dots, X_n)$  and  $Y = (Y_1, Y_2, \dots, Y_n)$  be two random vectors, and  $(x_1, x_2, \dots, x_n)$  and  $(y_1, y_2, \dots, y_n)$  be two particular realizations of  $X$  and  $Y$  respectively. Assume that the sub- $\sigma$  fields induced by  $X_k$ ,  $k = 1, 2, \dots, n$  are identical and independent. Similarly assume that the sub- $\sigma$  fields induced by  $Y_k$ ,  $k = 1, 2, \dots, n$  are also identical and independent. Let  $x_{(1)}, x_{(2)}, \dots, x_{(n)}$  be the values of  $x_1, x_2, \dots, x_n$ , and  $y_{(1)}, y_{(2)}, \dots, y_{(n)}$  be the values of  $y_1, y_2, \dots, y_n$  arranged in ascending order.

For  $X$  and  $Y$  if the empirical probability distribution functions  $\phi_1(x)$  and  $\phi_2(y)$  are defined as in (5) and (6) respectively. Then, the Glivenko-Cantelli Lemma of order statistics states that the mathematical expectation of the empirical probability distributions would be given by the respective theoretical probability distributions.

$$\phi_1(x) = \begin{cases} 0 & x < x_{(1)} \\ (r-1)/n & x_{(r-1)} \leq x \leq x_{(r)} \\ 1 & x \geq x_{(n)} \end{cases} \dots \quad (5)$$

$$\phi_2(y) = \begin{cases} 0 & y < y_{(1)} \\ (r-1)/n & y_{(r-1)} \leq y \leq y_{(r)} \\ 1 & y \geq y_{(n)} \end{cases} \dots \quad (6)$$

Now, let  $X_k$  is random in the interval  $[a, b]$  and  $Y_k$  is random in the interval  $[b, c]$  so that  $P_1(a, x)$  and  $P_2(b, y)$  are the

probability distribution functions followed by  $X_k$  and  $Y_k$  respectively. Then in this case Glivenko-Cantelli Lemma gives

$$\left. \begin{aligned} E[\phi_1(x)] &= P_1(a, x), a \leq x \leq b, \text{ and} \\ E[\phi_2(y)] &= P_1(b, y), b \leq y \leq c \end{aligned} \right\} \dots \quad (7)$$

It can be observed that in equation (4) the membership values of  $[x_{(r)}, x_{(r+1)}]^{(r/n)}$ ,  $r = 1, 2, \dots, n-1$  look like empirical probability distribution function  $\phi_1(x)$  and the membership values of  $[y_{(n-r)}, y_{(n-r+1)}]^{(r/n)}$ ,  $r=1,2,\dots,n-1$  look like the values of empirical complementary probability distribution function or empirical survival function  $[1 - \phi_2(y)]$ .

Therefore, if  $A(x)$  is the membership function of an L-R fuzzy number  $A=[a, b, c]$ . We get from (7)

$$A(x) = \begin{cases} P_1(a, x), & a \leq x \leq b \\ 1 - P_2(b, x), & b \leq x \leq c \end{cases} \quad (8)$$

Thus it can be seen that  $P_1(x)$  can indeed be the *Dubois-Prade* left reference function and  $(1 - P_2(x))$  can be the *Dubois-Prade* right reference function [14]. *Baruah* [13] has shown that if a possibility distribution is viewed in this way, two probability laws can, indeed, give rise to a possibility law.

#### IV. ALGORITHM PROPOSED

If the time-stamps stored in the transactions of temporal data are the time hierarchy of the type *hour\_day\_month\_year*, then we do not consider *year* in time hierarchy and only consider *day\_month*. We extract frequent itemsets using method discussed in [3]. Each frequent itemset will have a sequence of time intervals of the type  $[day\_month, day\_month]$  associated with it where it is frequent. Using the sequence of time intervals we can find the set of *superimposed* intervals [Definition of *superimposed* intervals is given in section-3] and each *superimposed* intervals will be a fuzzy intervals. The method is as follows: For a frequent itemset the set of *superimposed* intervals is initially empty, algorithm visits each intervals associated with the frequent itemset sequentially, if an interval is intersecting with the *core* of any existing *superimposed* intervals [Definition of *core* is given in section-3] in the set it will be *superimposed* on it and membership values will be adjusted else a new *superimposed* intervals will be started with the this interval. This process will be continued till the end of the sequence of time intervals. The process will

be repeated for all the frequent itemsets. Finally each frequent itemsets will have one or more superimposed time intervals. As the *superimposed* time intervals are used to generate fuzzy intervals, each frequent itemset will be associated with one or more fuzzy time intervals where it is frequent. Each *superimposed* intervals is represented in a compact manner discussed in section-3.

For representing each *superimposed* interval of the form  $[t^{(1)}, t^{(2)}]^{1/n} [t^{(2)}, t^{(3)}]^{2/n} [t^{(3)}, t^{(4)}]^{3/n} \dots [t^{(r)}, t^{(r+1)}]^{r/n} \dots$

$$[t^{(n)}, t^{(1)}]^{1/n} [t^{(1)}, t^{(2)}]^{2/n} \dots [t^{(n-2)}, t^{(n-1)}]^{n-2/n} [t^{(n-1)}, t^{(n)}]^{1/n}$$

we keep two arrays of real numbers, one for storing the values  $t^{(1)}, t^{(2)}, t^{(3)}, \dots, t^{(n)}$  and the other for storing the values  $t^{(1)}, t^{(2)}, \dots, t^{(n)}$  each of which is a sorted array. Now if a new interval  $[t, t']$  is to be *superimposed* on this interval we add  $t$  to the first array by finding its position (using binary search) in the first array so that it remains sorted. Similarly  $t'$  is added to the second array.

Data structure used for representing a *superimposed* interval is *struct superinterval*

```
{ int arsize, count;
  short *l, *r;
}
```

Here *arsize* represents the maximum size of the array used, *count* represents the number of intervals *superimposed*, and *l* and *r* are two pointer pointing to the two associated arrays..

#### A. Algorithm

for each locally frequent item set *s* do

```
{L ← sequence of time intervals associated with s
  Ls ← set of superimposed intervals initially set to null
  lt = L.get();
  // lt is now pointing to the first interval in L
  Ls.append(lt);
  while ((lt = L.get()) != null)
  {flag = 0;
    while ((l = Ls.get()) != null)
    {if(compsuperimp(lt, l))
      flag = 1;
      if (flag == 0) Ls.append(lt);
    }
  }
```

compsuperimp(lt, lst)

```
{ if( ! intersect(lst, lt) ) != null
```

```
{ superimp(lt, lst);
  return 1;
}
return 0;
}
```

The function *compsuperimp(lt, lst)* first computes the intersection of *lt* with the *core* of *lst*. If the intersection non-empty it superimposes *lt* by calling the function *superimp(lt, lst)* which actually carries on the *superimposition* process by updating the two lists associated as described earlier. The function returns 1 if *lt* has been *superimposed* on the *lst* otherwise returns 0. *get* and *append* are functions operating on lists to get a pointer to the next element in a list and to append an element into a list.

#### B. Estimate of the work done

Let *n* be the size of sequence of time intervals associated with a frequent itemset. Let *m* be the average number of intervals *superimposed* in one place. For each time interval of an itemset a pass is made through the list of *superimposed* intervals to check whether it can be *superimposed* on any of the existing *superimposed* intervals or not. Here each *superimposed* interval can generate fuzzy interval as shown in section-3. For this the intersection of the *core* of the superimposed time intervals and the current time interval is to be computed and this require  $O(1)$  time. If the interval *superimposes* then the time boundaries are to be inserted in the two-sorted arrays maintained for the superimposed intervals. Searching in a sorted array of size *m* requires  $O(\log m)$  time, inserting it in the current place requires  $O(m)$  time. Two such insertions will take  $O(2m)$  i.e.  $O(m)$  time. Thus for one itemset this process will require  $O(npm)$  time where *p* is the size of the set of superimposed intervals. Now  $p = O(n)$  and  $m = O(n)$ , thus the overall complexity in the worst case is  $O(n^3)$ . This will have to be done for each frequent item sets.

## V RESULTS OBTAINED

For experimental purpose, we have used a synthetic dataset T10I4D100K, available from FIMI<sup>1</sup> website. A summarized view of the dataset describing the number of items, the number of transactions, and the minimum, maximum and average length of transactions is presented in table 2. Since the dataset is non-temporal it cannot be used in its current form for our experimentation. The dataset mentioned in table 1 and some obtained results are presented in table 2.

TABLE 1. T10I4D100K DATASET CHARACTERISTICS

t	Datase	#Ite ms	#Transactio ns	Min[T]	max[T]	Avg [T]
	T10I4 D100K	942	100 000	4	77	39

<sup>1</sup> <http://fimi.cs.helsinki.fi/data/>

TABLE2. YEARLY FUZZY FREQUENT ITEMSETS FOR DIFFERENT SET OF  
TRANSACTIONS FOR ITEMSET {1}

Data Size (No of Transactions)	No. fuzzy time intervals
10000	1
20000	2
30000	2
40000	3
50000	3
60000	4
70000	4
80000	4
90000	4
100000	4

Here we keep the life-span of our datasets as 5 years. Firstly, we take only 10,000 transactions and found that the itemset {1} has a superimposed intervals superimposed on one place and hence it has one fuzzy time interval where it is frequent. For 20,000 and 30,000 transactions the same itemset has two superimposed intervals and so two fuzzy intervals, Finally from 60,000-100,000 transactions, we get {1} is frequent in four fuzzy time intervals.

#### VI CONCLUSIONS AND LINES FOR FUTURE WORK

An algorithm for finding yearly fuzzy patterns is discussed in this paper. The method takes input as a list of time intervals associated with a frequent itemset. The frequent itemset is generated using a method similar to the method discussed [4]. However, in our work we do not consider the *year* in the time hierarchy and only consider *month* and *day*. So each frequent itemset will be associated with a sequence of time intervals of the form [day\_month, day\_month] where it is frequent. The algorithm visits each interval in the sequence one by one and stores the intervals in the *superimposed* form. This way each frequent itemset is associated with one or more *superimposed* time intervals. Each *superimposed* interval will generate a fuzzy time intervals. In this way we will have each frequent itemset is associated with one or more fuzzy time intervals. An example such yearly fuzzy pattern is *cold-drinks* is frequent in every *summer*. The nicety about the method is that the algorithm is less user-dependent i.e. fuzzy time intervals are extracted by algorithm automatically.

Future work may be possible in the following ways.

- Other type of fuzzy patterns namely monthly and Daily patterns can be extracted.
- Clustering of patterns can be done based on their fuzzy time interval associated with yearly patterns using some statistical measure.

#### REFERENCES

[1] R. Agrawal, T. Imielinski, and A. N. Swami; Mining association rules between sets of items in large databases, *In Proc. of 1993 ACM SIGMOD Int'l Conf*

on Management of Data, Vol. 22(2) of SIGMOD Records, ACM Press, (1993), pp 207-216.

[2] J. M. Ale, and G. H. Rossi; An Approach to Discovering Temporal Association Rules, *In Proc. of 2000 ACM symposium on Applied Computing* (2000).

[3] A. K. Mahanta, F. A. Mazarbhuiya, and H. K. Baruah; Finding Locally and Periodically Frequent Sets and Periodic Association Rules, *In Proc. of 1<sup>st</sup> Int'l Conf. on Pattern Recognition and Machine Intelligence, LNCS 3776* (2005), pp. 576-582.

[4] A. K. Mahanta, F. A. Mazarbhuiya, and H. K. Baruah (2008). Finding Calendar-based Periodic Patterns, *Pattern Recognition Letters, Vol. 29(9), Elsevier publication, USA*, pp. 1274-1284

[5] C. M. Antunes, and A. L. Oliviera; Temporal Data Mining an overview, *Workshop on Temporal Data Mining-7<sup>th</sup> ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining*, (2001).

[6] B. Ozden, S. Ramaswamy, and A. Silberschatz; Cyclic Association Rules, *In Proc. of the 14<sup>th</sup> Int'l Conf. on Data Engineering, USA* (1998), pp. 412-421.

[7] Y. Li, P. Ning, X. S. Wang, and S. Jajodia; Discovering Calendar-based Temporal Association Rules, *Elsevier Science*, (2001).

[8] G. Zimbrado, J. Moreira de Souza, V. Teixeira de Almeida, and W. Araujo de Silva; An Algorithm to Discover Calendar-based Temporal Association Rules with Item's Lifespan Restriction, *In Proc. of the 8<sup>th</sup> ACM SIGKDD 2002*.

[9] R.B.V. Subramanyam, A. Goswami, Bhanu Prasad; Mining fuzzy temporal patterns from process instances with weighted temporal graphs, *Int. J. of Data Analysis Techniques and Strategies*, 2008 Vol.1, No.1, pp.60 – 77.

[10] S. Jain, S. Jain, and A. Jain; An assessment of Fuzzy Temporal Rule Mining, *International Journal of Application or Innovation in Engineering and Management (IJAEM)*, Vol. 2, 1, January 2013, pp. 42-45.

[11] Wan-Ju Lee, Jung-Yi Jiang and Shie-Jue Lee; Mining fuzzy periodic association rules, *Data & Knowledge Engineering*, Vol. 65, Issue 3, June 2008, pp. 442-462.

[12] Klir, J. and Yuan, B.; Fuzzy Sets and Logic Theory and Application, *Prentice Hill Pvt. Ltd.* (2002).

[13] H. K. Baruah; Set Superimposition and its application to the Theory of Fuzzy Sets, *Journal of Assam Science Society*, Vol. 10 No. 1 and 2, (1999), pp. 25-31.

[14] D. Dubois and H. Prade; Ranking fuzzy numbers in the setting of possibility theory, *Inf. Sc.*30, (1983), pp. 183-224.

## AUTHOR'S PROFILE



**Fokrul Alom Mazarbhuiya** received B.Sc. degree in Mathematics from Assam University, India and M.Sc. degree in Mathematics from Aligarh Muslim University, India. After this he obtained the Ph.D. degree in Computer Science from Gauhati University, India. He had been working as an Assistant Professor in College of Computer Science, King Khalid University, Abha, Kingdom of Saudi Arabia from October 2008 to September 2011. Currently he is serving as Assistant Professor at College of Computer Science and IT, Albaha University, Kingdom of Saudi Arabia. His research interest includes Data Mining, Information security, Fuzzy Mathematics and Fuzzy logic.

# Assessment of Non Transmittable Codewords Enhancement to Viterbi Algorithm Decoding

Salehe I. Mrutu<sup>1</sup>, Anael Sam<sup>2</sup> and Nerey H. Mvungi<sup>3</sup>

<sup>1,2</sup> School of Computational and Communication Science and Engineering (CoCSE),

Nelson Mandela African Institution of Science and Technology (NM-AIST), Arusha, Tanzania

<sup>3</sup> College of Information and Communication Technologies, University of Dar Es Salaam (UDSM), Dar Es Salaam, Tanzania

**Abstract**—Researchers have shown that practical mobile communication channels introduce errors that are concentrated in a certain locality rather than random errors. These are burst errors caused by deep fading of the wireless channel or a lightning strike. The existing Viterbi Algorithm (VA) capable of correcting random errors is inefficient in correcting burst errors and therefore resulting in unacceptable amount of residual errors. This paper presents an assessment of Non-Transmittable Codewords (NTCs) enhancement technique to VA in decoding the received signals subjected to burst errors that may occur in poor channels. A hard decision, 1/2 rate and constraint length K is equal to 3 Viterbi Algorithm decoding technique, Binary Phase-Shift Keying (BPSK) and Additional White Gaussian Noise (AWGN) are components used in MATLAB software based simulation when assessing the proposed technique. Applying 6NTCs to VA decoder enables the decoder to reduce 83.7 percent of its residual errors. However, the technique reduces the encoder's data transmission rate from 1/2 to 1/6.

**Keywords**—Locked Convolutional encoder; Bust errors; Residual errors; Non-Transmittable Codewords (NTCs); Viterbi Algorithm Decoding; Data Interleaver

## I. INTRODUCTION

A pair of binary convolutional encoder and Viterbi decoder is one of the mostly used components in digital communication to achieve low error rate data transmission. Convolution codes are popular Forward Error Correction (FEC) codes in use today. This type of code was first introduced by Elias in 1955 [1], [2]. VA introduced in 1967 [3], is known to be the maximum likelihood decoding algorithm of Convolutional codewords transmitted over unreliable channel [4]. VA is efficient in decoding random errors that occurred in a channel. However, the occurrence of burst errors in a received data block results in uncorrected or residual errors. Practical mobile communication channels are sometimes affected by errors which are concentrated in a certain locality rather than random errors [5]. These burst errors occur due to deep fading of the wireless channel or a strike of lightning in case of poor weather conditions or

intensive interference with other radio communication systems in the environments [6].

VA decoder can increase its error correction capability by increasing its constraints length (its memory size) [7]. However, increasing the memory size beyond 10 leads the decoder into prohibitive delay(not preferred by most real time applications) due to exponential growth of its decoding computation complexity [8], [2]. For decades, VA had been dealing with burst errors using an interleaving utility support. Without an interleaver, burst errors drive a viterbi decoder's decision unit into a confusion state which leads the decoder into failure and thus resulting in residual errors [9].

The basic idea behind the application of interleaved codes is to shuffle the received data. This action leads to randomization of the received burst errors that are closely located and then apply the VA decoder. Thus, the main function done by interleaver at transmitter is to change the input symbol sequence. At the receiver, de-interleaver changes the received sequence to get back the original sequence as the one at transmitter. There are two main categories of Interleaver utilities in communication systems that are block and convolutional interleavers.

Block interleaver just writes the received data row by row in a matrix and read them out for transmission column by column. Fig. 1 demonstrates how the block interleaver and de-interleaver work to jumble the received data and disperse burst errors. If a sufficient interleaver depth (number of rows in the interleaver/de-interleaver) is applied, then Interleaver successfully removes the effects of burst errors and turns them into controllable pattern of random errors by VA decoder.

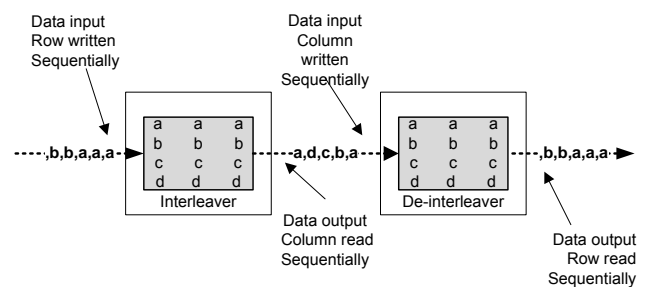


Fig. 1. Block Interleaver-De-interleaver





between a pair of received codeword from a channel and allowed codewords from the decoder at that particular time interval.

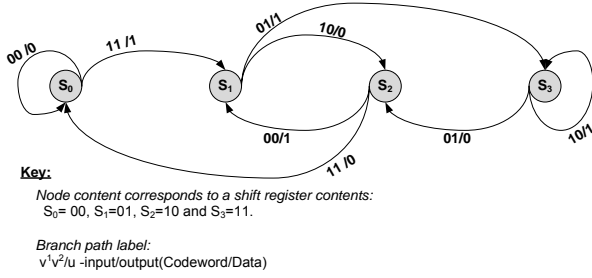


Fig. 3 Allowed state transition diagram of 1/2 decoding rate and Constraint length K=3 Viterbi Algorithm decoder

There are two ways of calculating HD [8] to find codeword's bits similarities and differences. In this paper codewords' bits similarities method is used. Therefore, similar bits are granted a value (i.e.1 HD) and non-similar bits have a zero value (i.e. 0 HD). In this case, results in each pair of comparison can be zero, one or two HDs. Fig.4 shows the calculation of HDs of each branch in each time interval and results are put in round brackets (i.e. (x)). After obtaining HD the algorithm continues as follows:

- Using a relation in (1) to recursively calculate Cumulative Branch Metrics (CBM) in each time interval  $t$  by adding the obtained  $HD_{(t)}$  to the  $CBM_{(t-1)}$  in each path and put results in a square bracket (i.e. [x]). Note that, for the time interval  $t=1$  there is no  $CBM_{(t-1)}$ , thus its value is zero.

$$CBM_{(t)} = CBM_{(t-1)} + HD_{(t)} \quad (1)$$

- At each node, find the path having the highest CBM up to time  $t$  by comparing CBMs of all paths converging to that node. In this step, decisions are used to recursively update the survivor path of that node. Equation (2) shows how the survivor path is obtained.

$$PM_{(t)} = \max(CBM_{(t)}^1, CBM_{(t)}^2) \quad (2)$$

- Eventually when the decoder terminates at time interval  $t$ , survivor paths leading to each node are compared to obtain a Winning Survivor Path (WSP). Equation (3) shows this relation. If more than one node has the same highest WPM then one of them is randomly selected and data are extracted from it.

$$WPM_{(t)} = \max(PM_{(t)}^1, PM_{(t)}^2, PM_{(t)}^3, PM_{(t)}^4) \quad (3)$$

Figure 4 is a trellis diagram of 1/2 decoding rate and constraint length K=3 Viterbi decoder demonstrating the discussed steps. The same codewords obtained in subsection A of this part {i.e. 11-01-01-00...} are assumed to have been

received with a bit error in the second bit of its third codeword {i.e. 11-01-00-00...}.

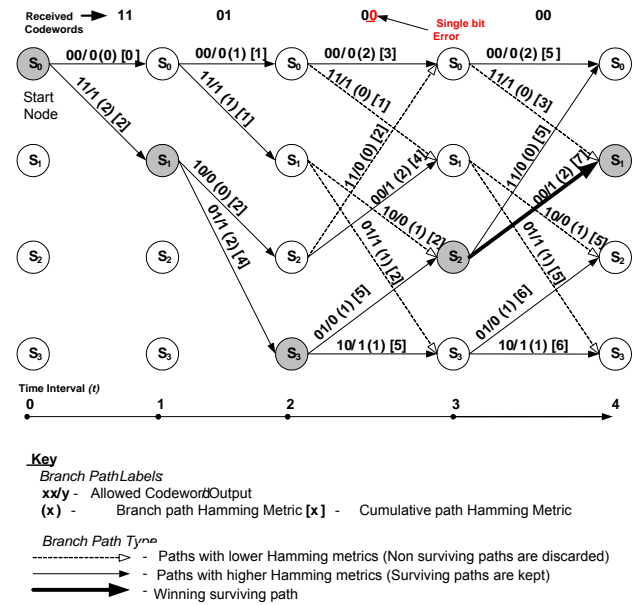


Fig. 4 Trellis diagram of 1/2 decoding rate and Constraint length K=3 Viterbi Algorithm decoder

### III. NON TRANSMITTABLE CODEWORDS ENHANCEMENT

Non transmittable Codewords (NTCs) technique can be applied at the data receiving machine where data encoded by a locked convolutional encoder arrives to be decoded by a VA decoder [9]. There are two different ways of locking a 1/2 rate and constraint length K=3 convolutional encoder. The methods include either adding two zero bits (i.e. 00) to the encoder after every data bit to be encoded (for the lower end locked encoder) or adding two one bits (i.e. 11) after each data input bit (for the higher end locked encoder) [9]. All examples and simulation in this paper applies a lower end locked 1/2 rate and constraint length K=3 convolutional encoder. Fig 5 shows the locking process. Lock bits reset a decoder to all zeros state after every data input bit. Suppose we have {1, 1...} as binary data stream ready for the encoding process. Fig. 5 shows the encoder locking process where letter "D" stands for a data bit or bits, and "L" stands for the integrated lock bits.

$$\overbrace{\{1-1...\}^D} \xrightarrow{\text{locking}} \{ \overset{D}{1} - \overset{L}{0} - \overset{L}{0} - \overset{D}{1} - \overset{L}{0} - \overset{L}{0} - \overset{L}{0} \dots \}$$

Fig.5. 1/2 rate and K=3 Convolutional encoder locking Process

After the encoding process, all codewords corresponding to both data and lock bits are transmitted over a noisy channel to the receiving machine. This fact lowers the data transmission rate of the encoder from 1/2 to 1/6. NTCs are known all zero codewords for the lower locked convolutional encoders that are added to the received codewords to channel to enhance the VA decoder in correcting the received errors

$$\{\overbrace{11}^D - \overbrace{10}^L - \overbrace{11}^L - \overbrace{11}^D - \overbrace{10}^L - \overbrace{11}^L \dots\} \xrightarrow{\text{locked}+2\text{NTCs}} \{\overbrace{00}^N - \overbrace{00}^N - \overbrace{11}^D - \overbrace{10}^L - \overbrace{11}^L - \overbrace{00}^N - \overbrace{00}^N - \overbrace{11}^D - \overbrace{10}^L - \overbrace{11}^L \dots\}$$

Fig.6. 1/2 rate and K=3 Convolutional encoder locking and 2NTCs addition process

successfully. NTCs can be added as a one, two, three codewords and so on; to each codeword corresponding to data bit. Suppose an example in fig. 5 {i.e. 1-0-0-1-0-0...} were encoded using the procedure discussed in table 1, the following codeword stream {11-10-11-11-10-11...} could be obtained for transmission. Fig. 6 demonstrates how 2NTCs are integrated to the received codewords corresponding to both data and lock bits before the received codewords are submitted for decoding. After the decoding process, all bits corresponding to the received lock codewords and the added NTCs are removed and the remaining data are submitted for use. In fig. 6 a letter “D” indicates a codeword corresponding to data bit, letter “L” is a codeword corresponding to lock bit and “N” is the added NTC.

#### IV. MODEL DESCRIPTION AND SIMULATION

This work, evaluates error correction capability of VA decoder and the Enhanced VA (EVA) decoder supported by NTCs. Both VA and EVA decoders, decode codewords from the same 1/2 rate and constraint length K=3 binary convolutional encoder. However, the encoder is locked using the discussed technique for the EVA decoder to enable it to utilize the technique described in section III of this paper. The number of residual errors from both the decoders forms a performance comparison factor between the two decoders. Therefore, a decoder with less residual errors is identified.

The Performance Measure of error correcting capability of the implemented codes is also given by Bit Error Rate (BER), which is obtained by the number of erroneous bits divided by the total number of transmitted bits. BER is affected by several factors including quantization technique used, noise in the channel, energy per symbol to noise ratio ( $E_s/N_o$ ), code rate, and transmitter power level [13]. In their work [14], Akyildiz and colleagues showed that BER is directly proportional to the code rate and inversely proportional to energy per symbol noise ratio and transmitter power level. The use of a proper decoder that corrects errors controls the increase in BER in transmitted data. The difference in BER that can be achieved by using error correction codes to that of uncoded transmission is known as coding gain.

A MATLAB software simulation that follows the procedures described in a block diagram described in fig. 7 performs the following:

- Generation of random binary data (i.e. 0 and 1) ;
- Addition and removal of encoder lock bits and NTCs for the case of EVA;
- Encode binary data using rate 1/2, generator polynomial  $[7,5]_8$  Convolutional code;
- Passing codewords through a noisy channel;
- Modulate and demodulate the codeword signals using hard decision technique for decoding process;

- Pass the received coded signals to Viterbi decoder and Enhanced Viterbi decoder;
- Counting the number of residual errors from the output of Viterbi decoder and Enhanced Viterbi decoder; and
- Repeating the same for multiple Signal-to-Noise Ratio (SNR) values

All the comparisons assume that both the algorithms have the same execution time. Table 2 shows the list of all parameters chosen for simulation.

TABLE II. SIMULATION PARAMETERS

Parameter	Value
Data length	$10^6$
Constraint Length (K)	3
Generator polynomial	$(7,5)_8$
Rate (r)	1/2
Encoder lock bits	2 zero bits (i.e. 00)
NTCs	1,2,3,4,5,6,7,8,9,10,11,12
Modulation/Demodulation	BPSK
Noise model	AWGN
Quantization	Hard Decision
Path evaluation	Hamming Distance Metric

##### A. Implementation of Codes

Figure 7 illustrates the procedure of encoding and decoding in a communication system, where randomly binary generated data from binary data source are sent directly to the binary convolutional encoder and data that will be decoded by EVA are sent through the lock bit addition node before they are submitted to the encoder. Codewords from the encoder are submitted to the discrete channel.

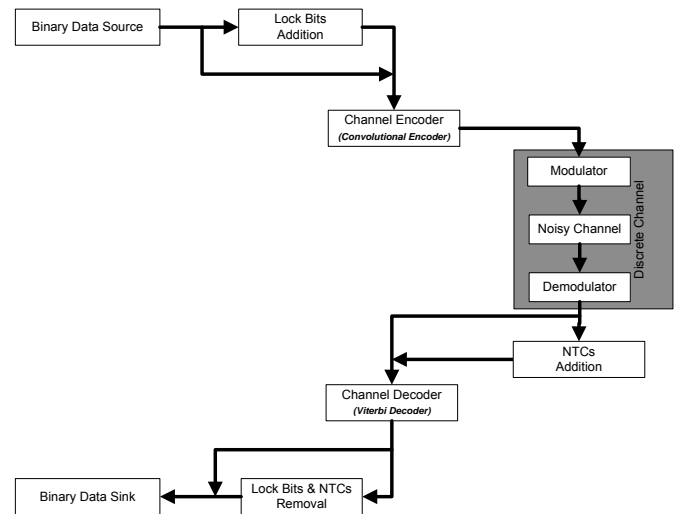


Fig. 7 Binary communication system block diagram used in simulation

The discrete channel modulates and demodulates the sent signal using Binary Phase-Shift Keying (BPSK) where zero bit (i.e. 0) is mapped to (-1) and one bit (i.e. 1) is mapped to (+1) and back, after the modulation signals are then released to the Additive White Gaussian Noise (AWGN) channel. Adding AWGN to signals in the transmission channel involves generating Gaussian random numbers, scaling the numbers according to the desired energy per symbol to noise density ratio ( $E_s/N_0$ ), and adding the scaled Gaussian random numbers to the channel symbol values.

Received codewords from the discrete channel are either sent to VA decoder directly or routed through NTCs node for adding NTCs. After decoding both lock bits and bits corresponding to the added NTCs are removed from data stream. Eventually, data obtained from the two streams (VA and EVA) are compared with the original generated data to determine the number of residual errors in each SNR value.

### B. Performance Measure

The theoretical uncoded BER using a relation in (4) is used to compare the code gain [7]. Where,  $E_b/N_0$  is expressed as a ratio of the involved factors; and “*erfc*” is a complementary error function in MATLAB software. For uncoded channel,  $E_s/N_0 = E_b/N_0$ , since there is one channel symbol per bit.

$$BER = 0.5 * erfc(\sqrt{E_b / N_0}) \quad (4)$$

However, the coded channel uses a relation (5) in the simulation.

$$E_s / N_0 = E_b / N_0 - 10 \log_{10}(2) \quad (5)$$

## V. RESULTS AND DISCUSSION

This section presents performance comparison between the VA and EVA decoders basing on their error correction capability in terms of BER and the residual errors. Fig. 8 compares the BER of uncoded channel, VA and 6 NTCs-EVA. Table 3 presents the counted residual errors from both VA and 6NTCs-EVA and the improvement obtained in each SNR value.

### A. Code gain

It is clear from fig. 8 that 6NTC-EVA decoder has the lowest BER curve with the highest constant code gain of 2 dB almost over all SNR values. While the VA decoder has highest BER (above theory uncoded) below 4dB and it is persistently higher than that of 6NTC-EVA. The minimal VA code gain is minus two (-2) dB. It is important to note that, around and below 4 dB, VA has a negative code gains because VA faces difficult in decoding burst errors in this area. However, 6NTCs-EVA performs better than both VA and the theory-uncoded curves with a difference of more than 2dB. It can also be observed that, The VA and 6NTCs-EVA curves are far apart in lower SNR values (let say below 4 dB)

and tend to come closer and closer as SNR values increase. This is because there are more errors generated in low SNR values which results in burst errors and create a great challenge to VA decoder. As the SNR value increases, few and random errors are generated and therefore VA decoder gains error correction power.

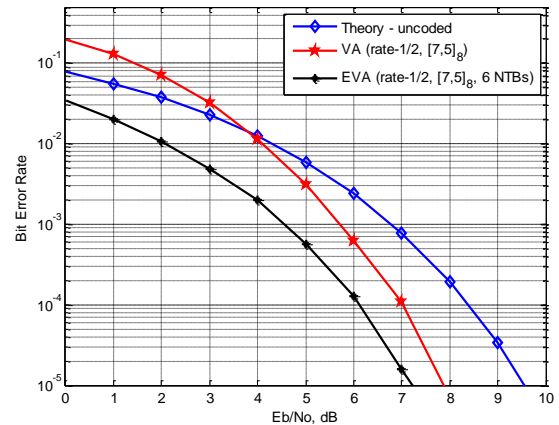


Fig.. 8. BER performance for theory-uncoded; VA and EVA in BPSK and AWGN

### B. Residual Error

Table 2. Compares residual error obtained from the simulation between VA and 6NTCs-EVA. The results show that, 83.7 percent of total residual errors occurred in VA were corrected by applying 6NTC to the EVA. Averagely, 84.3 percent of residual errors that occurred in and below 4 dB in VA were successfully corrected by 6NTC-EVA decoder

TABLE III. VA VERSUS 6NTCs-EVA RESIDUAL ERRORS

Eb/No, dB	VA Residual Errors	6NTCs-EVA Residual Errors	Data Error Recovery Improvement (Bits)	Data Error Recovery Improvement (Percentage)
1	198187	34407	163780	82.6
2	129604	20417	109187	84.2
3	72308	10650	61658	85.3
4	32492	4824	27668	85.2
5	11581	1985	9596	82.9
6	3094	571	2523	81.5
7	614	127	487	79.3
8	110	16	94	85.5
9	7	2	5	71.4
10	0	0	0	0.0
<b>Total</b>	<b>447997</b>	<b>72999</b>	<b>374998</b>	<b>83.7</b>

### C. Impact of Various NTCs Values on EVA

NTCs can be added as one, two, three codewords and so on. Fig. 9 shows that the increase in number of NTCs to EVA has an increasing impact on the decreasing rate of the number of residual errors. It further shows that, there is no significant reduction of residual errors with the further increase in the number of NTCs after 6 NTCs. These results concur with the explanation given by researchers in their work [9]

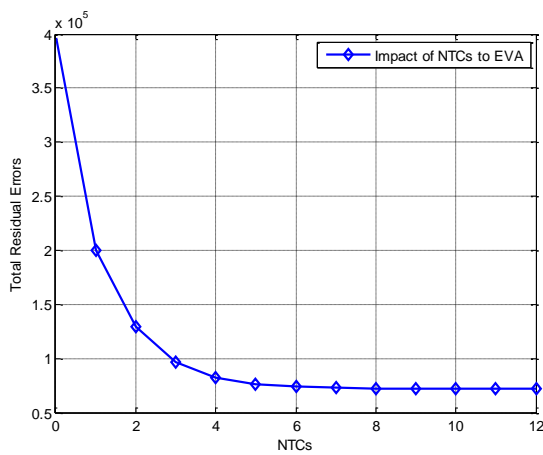


Fig. 9 Impact of NTCs to EVA

## VI. CONCLUSIONS & RECOMMENDATIONS

This paper presented and assessed the NTCs-enhancement technique to Viterbi Algorithm at the receiving machine. The

### REFERENCES

- [1] R. H. Morelos-Zaragoza, *The Art of Error Correcting Coding*, Second ed. The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.: John Wiley & Sons Ltd., 2006.
- [2] A. Viterbi, "Convolutional Codes and Their Performance in Communication Systems," *Communication Technology, IEEE Transactions on*, vol. 19, pp. 751-772, 1971.
- [3] A. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *Information Theory, IEEE Transactions on*, vol. 13, pp. 260-269, 1967.
- [4] D. Forney Jr, "Convolutional Codes II. Maximum-Likelihood Decoding," *Information and control*, vol. 25, pp. 222-266, 1974.
- [5] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2005.
- [6] Y. Zhang, R. Yu, S. Xie, W. Yao, Y. Xiao, and M. Guizani, "Home M2m Networks: Architectures, Standards, and QOS Improvement," *Communications Magazine, IEEE*, vol. 49, pp. 44-52, 2011.
- [7] C. Fleming. (2006, August, 13). *A Tutorial on Convolutional Coding with Viterbi Decoding. Spectrum Applications*. Available: <http://home.netcom.com/~chip.f/viterbi/tutorial.html>
- [8] S. I. Mrutu, A. Sam, and N. H. Mvungi, "Forward Error Correction Convolutional Codes for RTAs' Networks: An Overview," *International Journal of Computer Network and Information Security (IJCNIS)*, vol. 6, pp. 19-27, 2014.
- [9] S. I. Mrutu, A. Sam, and N. H. Mvungi, "Trellis Analysis of Transmission Burst Errors in Viterbi Decoding " *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 12-8, pp.46-53, 2014.
- [10] P. Gaj, J. Jasperneite, and M. Felser, "Computer Communication within Industrial Distributed Environment—a Survey," *Industrial Informatics, IEEE Transactions on*, vol. 9, pp. 182-189, 2013.
- [11] J. Silvestre-Blanes, L. Almeida, R. Marau, and P. Pedreiras, "Online Qos Management for Multimedia Real-Time Transmission in Industrial Networks," *Industrial Electronics, IEEE Transactions on*, vol. 58, pp. 1061-1071, 2011.
- [12] Z. Xu, S. Guan, and F. Yao, "A Novel Low-Time-Delay Convolutional Interleaver and Its Performance," in *Information and Communications Technologies (IETICT 2013), IET International Conference on*, 2013, pp. 208-212.
- [13] G. Balakrishnan, M. Yang, Y. Jiang, and Y. Kim, "Performance Analysis of Error Control Codes for Wireless Sensor Networks," in *Information Technology, 2007. ITNG'07. Fourth International Conference on*, 2007, pp. 876-879.

EVA is used to recover distorted codewords from a noisy channel. The decoder at the receiving machine recovers erroneous received codewords from a 1/2 rate, constraint length K is equal to 3 convolutional encoder at the sender point. MATLAB software was designed where a hard decision modulation and demodulation schemes using Binary Phase-Shift Keying (BPSK), Addition White Gaussian Noise (AWGN) were implemented. The simulation results showed 83.7 percent overall improvements in reducing residual errors when 6 NTCs were applied to VA decoder. This is a very significant improvement to VA decoders. The enhanced VA can be used in industries that demand for error free transmission such as telemedicine. However, the technique lowered the encoder's data transmission rate from 1/2 to 1/6. Further research of the proposed technique is highly recommended to show the impact of the technique in different platforms and applications using Viterbi Algorithm.

- [14] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless Sensor Networks: A Survey," *Computer networks*, vol. 38, pp. 393-422, 2002.

### AUTHORS PROFILE



**Salehe I. Mrutu** received his B.Sc. and M.Sc. degrees in Computer Science from The International University of Africa in 2003 and the University of Gezira in 2006 respectively. He is currently a Ph.D. scholar at the school of Computational and Communication Science and Engineering of the Nelson Mandela African Institution of Science and Technology in Arusha, Tanzania. He is also serving as assistant lecturer at the University of Dodoma under the school of informatics since 2007. His research interests include Forward Error Correction codes, quality-of-service provisioning and resource management for multimedia communications networks.



**Anael Sam** received his B.Sc., M.Sc. and Ph.D. in Electronics Engineering (Institute of Electronics and Photonics, Slovak University of Technology, Slovak Republic). He works as senior lecturer at the Nelson Mandela Institution of Science and Technology, Arusha, Tanzania. Dr Sam's specialization and research interests are in radio, multimedia and mobile communication systems; electronics and telecommunication engineering, software quality assurance engineering and mobile networks optimization. He is also a member of IEEE and ISQTB international.



**Nerey H. Mvungi** received the B.Sc. degree in electrical engineering from the University of Dar Es Salaam, Tanzania, in 1978; the M.Sc. degree in electronics control from Salford University, U.K. in 1980; and the Ph.D. degree from Leeds University Leeds, U.K. in 1989. He worked for a year with the Phillips Center for Technology, Eindhoven, Eindhoven, and the Netherlands February 1992 to Feb 1993. He was attached to Onersol Solar Energy Research Centre in Niamey June-July 1991 as ILO Consultant on Solar Energy Systems. Since his undergraduate graduation in 1978, he has worked as an academician and is now a full professor. He has mostly worked in the University of Dar es Salaam but for the period of September 2008 to June 2012 when he was at the University Dodoma in Tanzania to starting a new IT College as its founding Principal. Prof. Mvungi's research interests are in control and instrumentation, computer communication and applied electronics, lightning protection, rural access, power-quality aspects, and remote monitoring and control of energy consumption and digital broadcasting. He received a 2010 IBM Faculty Award.

# An Integrated Digital Academic Repository Model for Higher Learning Institutions (Case of Tanzania)

Martha Mhongole\*

School of computational communication science and  
engineering  
NM-AIST, Arusha, Tanzania

Loserian Laizer

School of computational communication science and  
engineering  
NM-AIST, Arusha, Tanzania

**Abstract**— This paper explores the current existing models and technologies used in knowledge creation, knowledge sharing and knowledge dissemination practices in Higher Learning Institutions (HLIs) of Tanzania and proposes the model for the development of an Integrated Digital Academic Repository that enhances management, sharing and dissemination of Scholarly works produced in HLIs of Tanzania. The proposed model is presented and described in the paper. The study was carried out in three HLI using questionnaires, interview, observation and review of literatures. The findings show that, universities produce wide range of intellectual outputs such as research articles, learning materials, theses and technical reports. More than half population involved in the study create and store their intellectual outputs in personal computer hard drives while others store in internet cloud servers and departmental web servers. Moreover, sharing and dissemination of Intellectual output is done through internet i.e. Emails, social network, institution website and cloud servers, journal publication, seminar presentations, posters and printed copies in libraries. The identified methods proven to be unreliable and hindering availability and accessibility of scholarly works. Thus the proposed model provide a central system through which intellectual outputs will be collected, organized and archived and disseminated through it. The paper concludes with the conceptual framework of the proposed system, whereas design and development carried forward to be our future work.

**Keywords**- Higher learning institution, intellectual output, knowledge management, knowledge sharing, model, digital repository

## I. INTRODUCTION

In today's world, knowledge has been considered as a strategy resources that formulate the knowledge-based economy of countries. Knowledge has been identified as important as other factors of production such as land, labor and capital that requires management for the development of society [1, 2]. Knowledge based economy is an economy in which knowledge is being created, acquired, transmitted and used more effectively by individuals, enterprises, organizations and communities to promote economic and social development[3,5]

Effective management, dissemination, sharing and use of knowledge assist in solving problems such as diseases, poverty, illiterate, environmental degradation and deforestation especially in African countries whose half population(50%) live in underprivileged societies, lacking access to information and suffer a lot when they fail to acquire and use information in their lives[3,4,5]. According to [6], every developed institution has a duty to place and disseminate knowledge through centers, which can easily be accessed in underprivileged society.

Higher Learning Institutions (HLIs) have been described as the canter of creativity, innovation and the main producers of knowledge, both scientific and technological, that need a reliable, technological, affordable and accessible media to manage and disseminate their scholarly work to the world [1]. Knowledge management is a practice of organizing, storing, and sharing of vital information, so that everyone can benefit from its use. Despite the number of practices in knowledge management process, knowledge sharing has been identified as the most important aspect of knowledge management process as it facilitate dissemination and application of the created knowledge [3]. As pointed out by [7], knowledge has to be shared and disseminated to the designated community for it to be used, otherwise it end in itself.

Researchers, student and faculty members in HLIs produce wide range of intellectual outputs such as research articles, datasets, theses, dissertation, reports, presentation and learning materials [8]. According to [5], Scholars work have to be available and accessible to Scholars and public in general. Scholars and Public should might apply the knowledge and use the disseminated information as a base for their research; this led to knowledge evolvment and economic development [4]. However, with materials produced being scattered allocated, dis organized and lack of systemic integrated system to

Despite, the rapid increase of digital materials produced in HLI [8], availability and discoverability of the scholarly work remained to be a challenge. About 80%-85% of HLIs outputs such as research articles, manuscripts particularly from African countries have never been made accessible and



discoverable to the scholarly community and the world [9, 10]. Literature Scholars archive their intellectual outputs onto their personal computer hard drives, departmental web servers and in library shelves of which access is guaranteed to limited number of people or none and also lost because most of the materials are not well organized and have no clear documentation [8,10]. Archaic dissemination techniques, subscription fee and publication charges have been pointed out among factors contributing to limited number of intellectual output and restrict accessibility over the scholarly works [4].

Unavailability and limited access to scholarly works present problems such as repetition of works done by other scholars, limit knowledge evolvement, waste national resource, effort and money as well as negatively affect countries development [11, 12].

To this end, this study explores the current situation of intellectual output sharing process, identifying the challenges, and propose a new model enhancing management, sharing and dissemination of scholarly works created in HLIs of Tanzania. In order to achieve this objective, the study is divided into the following specific objectives:

- 1) to identify types of intellectual output produced in Tanzanian HLI.
- 2) to analyze how scholarly works created in Tanzania HLI collected, organized, archived, managed, shared and communicated to scholarly community and the world
- 3) to identify the challenges associated with the current archiving and dissemination techniques of scholarly works used in Tanzanian HLI ;
- 4) to propose new model enhancing knowledge sharing and dissemination in HLI
- 5) to identify design requirements of a proposed model

## II. METHODOLOGY

The study was conducted in three HLIs named: Nelson Mandela African Institution of Science and Technology (NM-AIST), Muhimbili Health of Allied Sciences (MUHAS) and Sokoine University of Agriculture (SUA). The chosen study area are the science universities offering undergraduate and postgraduate studies, though for our study postgraduate students, researchers and faculty members were involved considering that, they at most produce intellectual output. The composition considered the importance of including main stakeholders producing and managing intellectual outputs in HLIs. Questionnaires and interview guide questions were used as data collection tools. Questionnaires were administered to researchers, students and faculty member's whereby, library managers were approached for face-to-face interviews. Library managers were interviewed for detailed information and experience on how their institutions manage and disseminate scholarly works. Questionnaires were designed to capture information on types of intellectual output produced and how they had been stored and disseminated to public. Respondents were also asked about the challenges associated with the current archiving and dissemination techniques. Detailed literature review on materials related to topic was

done to familiarize with existing digital contents sharing techniques and identifying challenges and weaknesses associated with each method .Data were summarized and analysed using the Statistical Package of Social Science (SPSS) and excel. Pictorial presentations of data were used to compare and derive important patterns that are to be used for further research.

## III. FINDINGS AND DISCUSSION

### A. Respondent profile

The respondent profile was meant to describe the respondent designation and educational level, from which authors were able to judge a respondent as appropriate individual or group who can create intellectual output, use, share and disseminate for other people to learn. A total of 95 questionnaires were administered to students, researchers and faculty members of the studied institutions. The population consisted of 67% students, 28% researchers and 5% faculty members as shown in Fig 1.

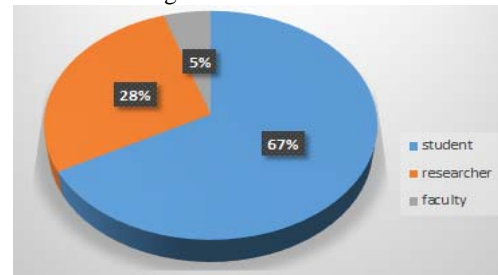


Fig 1 Population composition

Fig 2 shows response by level of education whereby 76% of respondents were master's students, 17% were PhD (candidates and holders) and 7% were administrative staff.

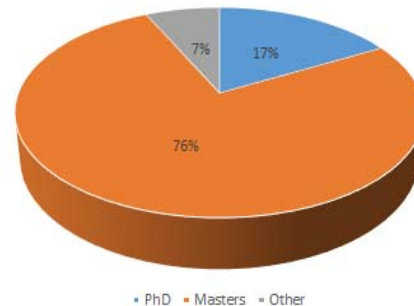


Fig 2 Respondents' level of education

We assumed the population is appropriate for production of intellectual outputs whereby, students may produce theses, dissertation, technical project reports, researchers come up with research findings and faculty members create learning materials such as lecture notes and presentations. We observed that, in HLI particularly for postgraduate studies, student must produce intellectual output such as research papers, theses and dissertation as criteria for graduation. Institution may take



advantage of the regulation by introducing a mandatory policy requiring scholars to self-archive and disseminate their research findings and learning materials through institutional database repository of which can be used as a source of scientific information to be used in research, academic or economic development. Moreover, a researcher can make use of the archived data as the base for next generation research development as well to display and publicize institution research product to the public. To achieve the goal, institutions need to establish a stable, permanent, accessible, affordable technology that will facilitate storage of large volume of intellectual outputs and dissemination to the large number of scholars and the public of which is a database.

### B. Intellectual output creation and Dissemination

The findings show that, Scholars in Tanzanian HLI are producing different types of intellectual outputs including leaning materials, research articles, manuscripts, technical project materials. Fig 3 shows that, 28% of respondents had engaged in production of learning materials such as presentation and class notes, 23% had created research articles, 21% technical reports, 21% created manuscripts and 7% did not specify type of intellectual outputs that they had ever created.

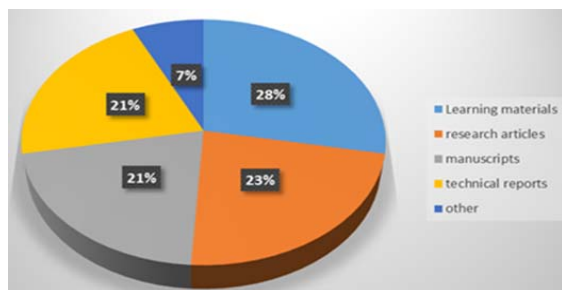


Fig 3 Intellectual output produced in HLI

The findings revealed that, abundantly knowledge and skills are being created by different scholars in HLI, which would be worth efficient to be preserved and disseminated to the designated community to be applied bot for academic and economic development. However, we usually come across research papers published by institutions on different journals, with no or little amount of other types of intellectual outputs produced by the same institutions. With this observation shows that, institutions does not considers other types of intellectual output as important as research articles. However, [7] suggested that, for material created to be of positive impact and useful ,one must collect, organize, archive, share and disseminate the disseminated and applied to the designated community to be used, otherwise it become a wastage of resources such as time ,money and effort engaging in production of objects knowing that they are not useful.

Personal computer hard drive, internet: emails, cloud servers and printed copies were mentioned as technologies used to archive and manage HLIs intellectual outputs. As

shown in Fig. 4. 51% of the respondents collect and archive their intellectual outputs in personal computer (PC) hard drive, 30% store on internet particularly on cloud servers such as email, Google drive and drop box, 17% print and preserve hard copies of their works and 3% were not certain about the methods they are using.

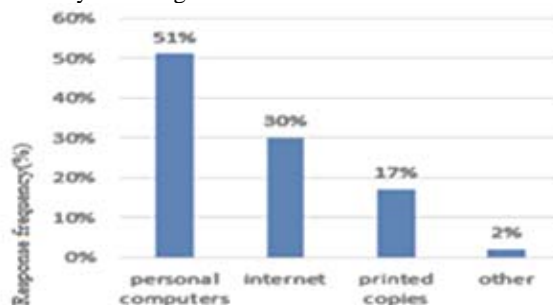


Fig 4 Intellectual output storage mechanisms

It has been realized that scholars in HLIs share and disseminate their research and academic works via different technologies such Internet: emails, social network, institution website and cloud servers, journal publication, seminar presentations, posters and printed copies in libraries. The result in Fig 5 shows that, 35% of respondents use internet as their content dissemination mechanism, 30% publish their output onto journals, 24% presents their outputs in seminar and workshops, 4% print and archive their copies, 4% publish their works through posters and 4% were not certain about the media they use. The findings revealed that, produced intellectual outputs are widely scattered stored and disseminated. From the findings, we observed that, searching and retrieving of contents, which are widely scattered, it consume time and use much more bandwidth compared to when the resources are retrieved from single source, that is well organized.

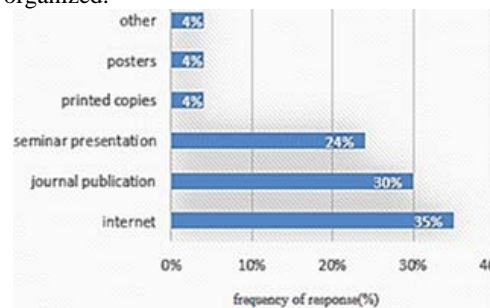


Fig 5 Intellectual output dissemination mechanism

Limited storage space, high publication cost, accessibility cost, limited internet connectivity, physical security, access and sharing limitation were mentioned as challenges in the in process of management and dissemination of intellectual outputs. As shown in Fig 6, 45% of respondents had ever experienced limited storage space to archive their output, 30% mentioned high publication and accessibility cost, 20% experienced limited access and sharing, 35% claimed

unreliability of their system (crash), and 10% identified physical security as their challenge.

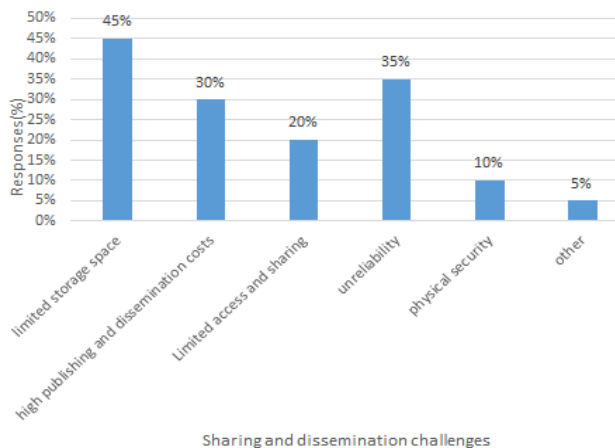


Fig 6 Sharing and dissemination challenges

#### IV. THE PROPOSED INTEGRATED DIGITAL ACADEMIC REPOSITORY MODEL FOR HLIS

The proposed model developed based on Open Archive Information System Reference Model (OAIS). The OAIS reference model is a conceptual framework for a generic archival system which is committed to a dual role of preserving and providing access to information. Central to the reference model is an open archival information system (OAIS) which is “an organization of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community” [14]. The model describes the functional components which collectively fulfil the system’s preservation and access responsibilities as shown in Fig 7.

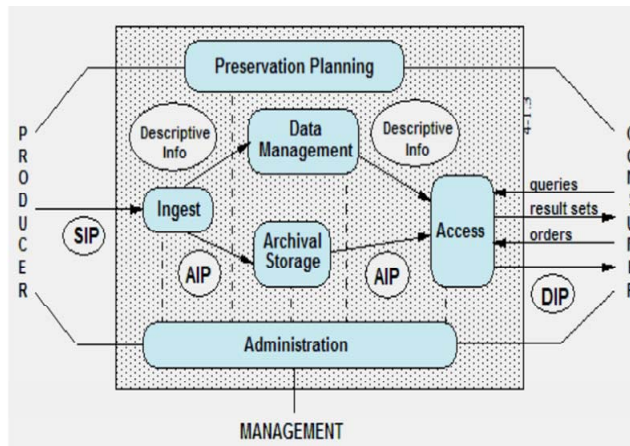


Fig 7 OAIS model

The functional components of an OAIS include:

- **Ingest** - services and functions that accept information submitted by Producers and prepare it for storage and management within the archive. In our case, Ingest involve the use of web user interface through which scholars (students, researchers and faculty) get access into a system and submit their contents of which need to be organized and stored in a database.
- **Archival storage**- manages the long-term storage and maintenance of the digital materials entrusted to the OAIS, to make sure they remain complete and render able over the long term. Media refreshment and format migration for example are typical procedures that would be undertaken by the archival storage function.
- **Data management**- maintains descriptive metadata to support search and retrieval of the archived content, and administration of internal operations. In our proposed model, content submitter is required to briefly describe the metadata/information that will be used to identify and retrieve the content from the database.
- **Preservation planning**- designs preservation strategy based on evolving user and technology environment
- **Access** -manages processes and services that locate, request, and receive delivery of the content within the archival store.
- **Administration** - responsible for day-to-day operations and the co-ordination of the five other OAIS services.

Having identified the challenges facing HLIs in managing, sharing and dissemination of intellectual outputs, we proposed a new model of intellectual output sharing and dissemination called integrated digital academic repository. The model assist the central management of intellectual output of HLI in a central database, whereby scholars from different universities in Tanzania will be able to put and access the materials from the repository. The proposed model will facilitate collection, archiving and dissemination of intellectual output which are created in Tanzania HLIs. Researchers, students and faculty member’s intellectual outputs will centrally be collected, reviewed, archived in and disseminated to the scholarly community and public. Integrated Digital academic repository system (IDAR), enables researchers to communicate research findings and find out what is been done by other researchers from their institutions and other universities on their field and other fields in HLIs of Tanzania. Scholars will have access to research results and learning materials of which can be used in academic or research activities as source of scientific information. Faculty members of various universities will be able to share their academic materials which are useful for research and academic purpose with scholars in HLIs and the globe over the internet as shown in Fig 8.

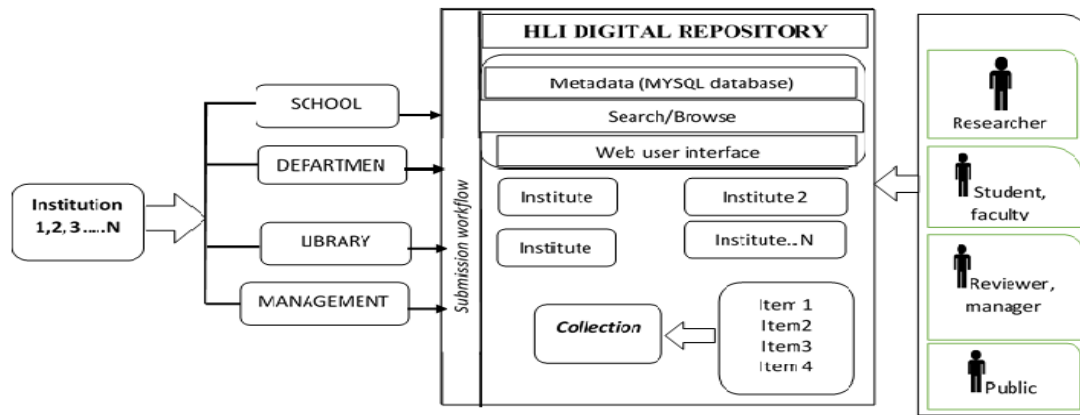


Fig 8 Proposed IDAR model

Fig 9 shows how the information flows from producers who for this case are students, researchers, and staff belonging to a particular HLIs. On web in the presence of internet, scholars from different universities and research centres submit their

IOs of which are organized and put together in a database (repository). Access is also initiated from the established repository.

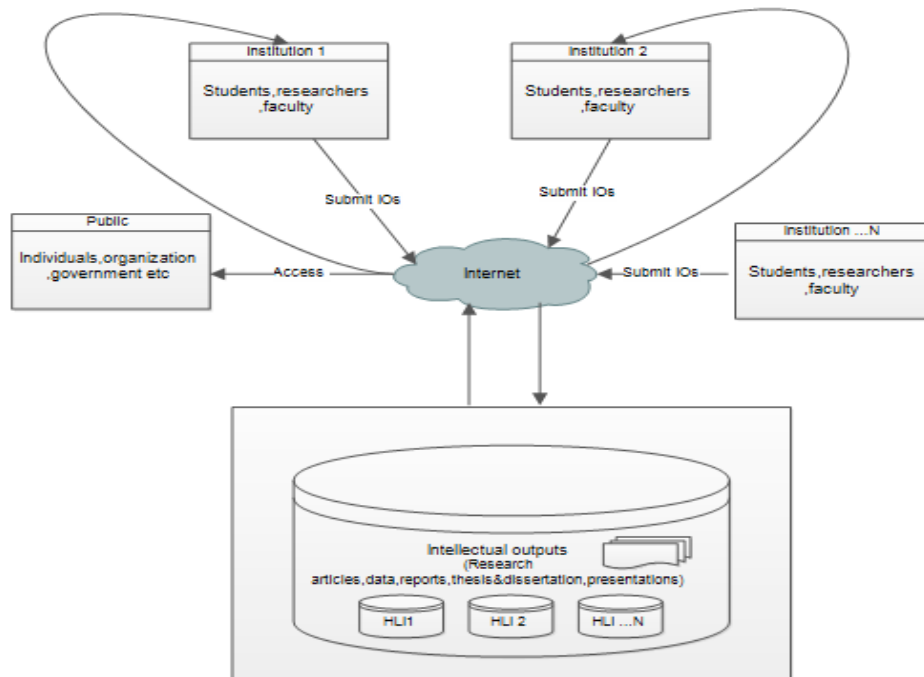


Fig 9 Conceptual IDAR Information Flow

#### A. User requirement specifications

This section presents user perspectives and requirements towards the development of the proposed system which were gathered during data collection. According to [15] requirements play vital important role and are a primary tool towards development of any information system. They define

what to be performed by a system and specify how it will be performed. Requirements are categorized into two groups namely: functional requirements and non-functional requirements. Functional requirements describe things, actions, tasks and functions that the system is required to perform or services the system should provide. Non-functional requirements describe properties and system constraints such as interface requirements, reliability, performance, storage

capacity, usability and system security. The non-functional requirement does not directly relate to the system functionalities though are the ones describing how the system function should be performed [16]. For the proposed model to be successfully implemented, the primary fundamental task was to identify the requirements defining the functional specifications to be incorporated into a design of the system. According to [17] for any information system to be successful developed, requirements must be gathered from different stakeholders and prospective users of the system. It has been observed that, as users getting involved in the system development process particular requirement elicitation, the possibility of developing a system which is usable and highly acceptable is high [18].

Therefore, we described and elaborated the proposed model to stakeholders (researchers, students and faculty), who in turn gave their inputs defining the type of the system needed, specifying system operations and organization of the contents that will be collected from scholars. Not only that but also users specified the mode of access to be provided to each user as follows:

*i. User Perspective towards Development of Integrated digital academic repository*

Despite the existence of digital repositories into some of the visited institutions: MUHAS and SUA that collects and disseminate scholarly works of their institutions and materials related to health and climate change respectively. Scholars of the stated universities joined hand with scholars at NM-AIST who were currently not possessing digital repository, supported the development of the proposed integrated digital academic repository as it expands and widens the search area of materials. The findings show that 97% of respondents supported the development of the proposed repository whereas, 3% did not support. Likewise the result shows that 87% of respondents were extremely interested, 10% were somewhat interested, 3% were neither interested nor uninterested and no one mentioned not to be interested. From the statistics, we realized that users are in need of the proposed system and it brought attention to us that we required to dig deeper for more information in order to get the requirement or services that users expect the system to provide

*ii. Type and format of the intellectual outputs*

From the study, respondents identified different types of intellectual outputs and different file formats to be archived in and disseminated through the proposed repository. Theses and dissertations, technical project reports, research articles and learning materials (lectures, seminar presentations) are the common items users demanded to be accommodated in a proposed system. Whereas, Text files (doc, pdf), Multimedia (video, audio) and Binary files were among the file formats identified by users that need to be uploaded and accessed from the system as shown in Fig 7.

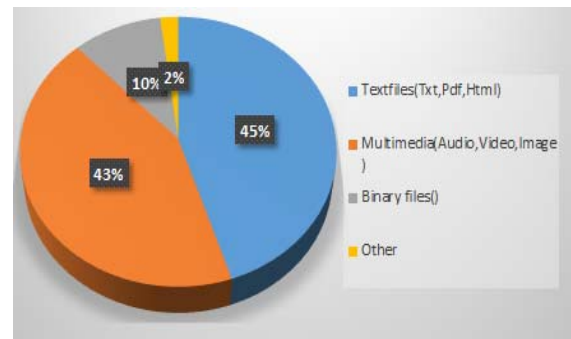


Fig 10 Intellectual output

*iii. Intellectual output submission and organization*

Majority (86%) of respondents showed their interest of accessing reliable information that has been proved by expertise as correct and useful, therefore reviewers have been proposed to check the submitted contents before uploading. Majority (86%) of respondents showed their interest of accessing reliable information that has been proved by expertise as correct and useful, therefore reviewers have been proposed to check the submitted contents before uploading to the system, 14% suggested the submitted materials to be uploaded directly by the corresponding submitter (researcher, student, faculty). From the study also, users suggested the way of organizing submitted digital contents whereby 41% of respondents demanded materials to be organized based on field of study, 26% authors name, 32% category wise (paper, books, articles, presentation) and 1% did not specify.

*iv. Intellectual output sharing and dissemination mode*

The result shows that, 60% of respondents need the archived intellectual outputs to be shared and disseminated to scholars within institution, outside institution and the public. Whereas, 33% requires the materials to be shared among the scholars within and outside institution, 3% prefer the materials to be shared only with scholars in institution and 3% were not certain about which mode to use. The fact that, not all produced materials are necessarily to be shared in globe, some are only necessary to a particular institution or people while others might be necessary to the globe. We considered all suggestion necessary and documented to be included in the further development stages, roles have been defined allowing user to specify whether the submitted materials should be available to all people accessing the system (public access) or institution members (institution) or to be archived and only accessed by the author/submitter of the work (private/individual access).

*B. System Functions*

This section summarizes list of functions to be performed by the proposed system following the requirements which were collected from users.



TABLE I. SUMMARY OF FUNCTIONAL REQUIREMENTS

No	Function
1	System must be able to register users(institutions and institution members)
2	System must allow user to submit and access contents
3	System must be able to identify system user using username and roles played by each user
4	System must provide a means through which submitted contents will be reviewed and edited before uploading for use
5	System must be able to generate reports for management on the status and trends of institution and user content contribution
6	The system should provide a way for content contributors to register themselves
7	System must be able to accommodate contents of various formats i.e. Text ,multimedia and binary files
8.	System must be able to send notification to users confirming their registration and status of their submission through emails

We used the use case diagram to present system functional requirements gathered from users. The Use Case Diagram is a Unified Modelling Language (UML) providing a pictorial representation of a system and how user interact with the system [19]. The use case diagram depicts the abstract view of the system as shown in Fig 11.

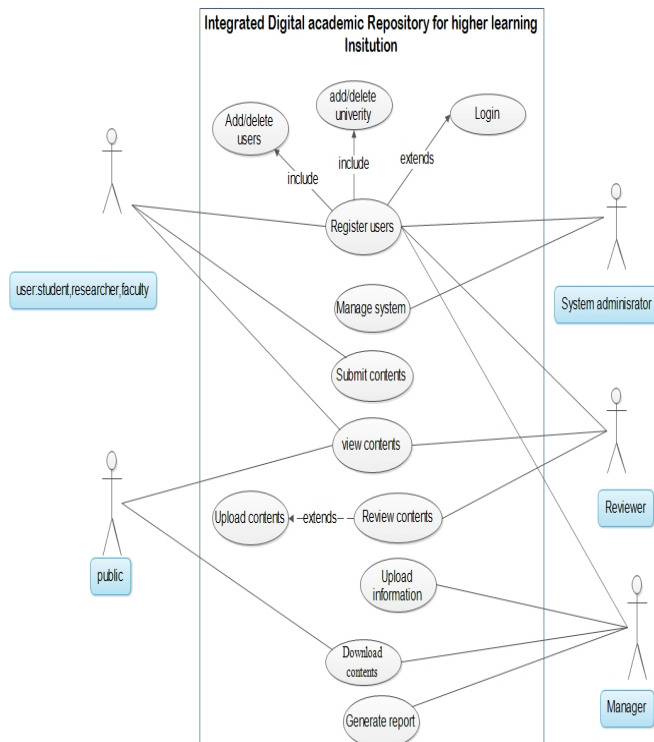


Fig 11 Use Case for IDAR

### C. Non-functional requirements

The fact that, scholars need access to materials produced by different scholars from different institutions which are in large volume of information and storage devices with high storage capacity is considered in the new proposed system. Materials collected and archived in repository in long term as a source of scientific information in future studies, back up techniques are necessary and useful in case of system crash and easy recovery. Web based interface will be provided for easy access to materials. Intellectual outputs will be organized into module basis following the institution, schools and departments on which a submitter belongs to minimize searching and retrieval time. Ensuring the reliability and usability of the archived contents, peer reviewing and editing are considered necessary in order to ensure the correctness and usefulness of the archived materials. Moreover, user IDs and roles have been defined to serve as security mechanism in the proposed system.

### D. Intellectual Property Right (IPR)

From literatures intellectual property rights (IPRs) such as copyright and license have been identified as the biggest road blocking to self-archiving that limits the number of intellectual outputs to be populated into a repository. The fact that intellectual outputs consist of innovative, creativity and skills of individuals who in turn likely to be recognized, possess and have control over their scholarly works, it happened that the materials are copyrighted or licensed to a particular person or group of people. The access of item that is described under particular IPRs depends on the terms and condition the law protecting the product of which follows under IPRs. IPRs are rights granted to creators and owners for their intellectual creativity in the industrial, scientific, literary, and artistic domain. The work can be in the form of an invention, a manuscript, a suite of software or a business name. The Rights have been introduced intentionally to protect content creators right and at the same time allowing the public to access their creativity [20]. It has been observed that, scholar's sacrificed invaluable right in expense of publishing their contents to publishers. However for self-archiving of copyrighted contents the law does not state directly on what should be done, instead it leaves decision to the parties involved in the publishing agreement. The agreement may be publisher allowing authors to self-archive copy of their published work into their repository but not for commercial gain. Also work published on Open Access (OA) journals termed as free materials to be access by anyone (green).Materials published and declared gold requires subscription or pay per-view fee whereby green materials and unpublished materials have no restriction on user access .Thus for our proposed system, we suggest to archive both the published and unpublished materials while preserving the IPRs of authors. The gold published materials will be archived with the main intent of alerting users of their presence, therefore the title and abstract metadata of the work will be provided as a pointer for further inquiries.

## V. CONCLUSION AND FUTURE WORK

The current situation of intellectual output sharing and dissemination in Higher Learning Institution of Tanzania has been identified and presented in this paper. Challenges and weaknesses facing sharing and accessing of academic and research works produced in Higher learning institution of Tanzania have been identified. It has been observed that, HLI produce large number of digital contents including research articles, theses, technical project reports, conference proceeding and learning materials. However access and availability of this materials remain to be a problem in scholarly community and the world. Universities lacks central system to collect, organize, manage and share their intellectual outputs. Materials are scattered allocated, published onto international journals of which access is limited by subscription fees and pay-per view fees, and some housed onto individuals hard drive or left unpublished in university library shelves providing of which access is granted to limited number of people or none around institution.

To address the identified challenges, the new intellectual output sharing model that enhances organization, accessing, sharing and dissemination of intellectual outputs which are created in Higher learning institution of Tanzania has been proposed and presented in this paper. The proposed model provide a platform that manages and disseminates created output in a central way. Functional and non-functional requirements of the proposed system have been identified and summarized. Therefore, using the proposed model and the identified user requirements, authors of this work have considered the design and development of the information system that portray the identified characteristics to be their future work.

## ACKNOWLEDGMENT

The authors would like to acknowledge the support provided by Nelson Mandela Africa Institute of Science and Technology for providing financial and material support for the research undertaking. They also recognizes the support provided by Sokoine University, Muhimbili university of Health and Allied Sciences and Nelson Mandela Africa Institution of Science and Technology by offering data collection venue and consultancy.

## REFERENCES

- [1] Córcoles, R.Y. (2013). Intellectual capital management and reporting in European higher education institutions. *Omniscience*. 9 (1), pp 2-16
- [2] Mavodza, J. and Ngulube, P., 2012, 'Knowledge management practices at an institution of higher learning', *SA Journal of Information Management* 14(1).
- [3] Oseghale O and Adeyomoye J. J. (2011). Emergence of the Global Knowledge Economy: Implications for libraries and lifelong learning in Nigeria. An online journal of the African Educational Research Network. 11 (2), pp 84-87
- [4] Chisenga, J. (2006). The development and use of digital libraries, institutional digital repositories and open access archives for research and national development in Africa: opportunities and challenges. WSIS

Implementation in Higher Education. Advanced Research in Scientific Areas-International Virtual Conference. Pp. 2078-2081

- [5] Pinto, J (2012). A Framework for Knowledge Management Systems Implementation in Higher Education. Advanced Research in Scientific Areas-International Virtual Conference. Pp. 2078-2081
- [6] Bakshi, R. (2013). Sharing and connecting knowledge in indian educational institutions through knowledge management. *International journal of behavioral social and movement sciences*. 3 (1), pp252-260.
- [7] Sivakumar, P. (2012). Knowledge Production & Dissemination: An Analysis in the Context of the National Youth Policy. *Journal of Management & Public Policy*. 4 (1), pp33-36
- [8] Pinto, J (2012). A Framework for Knowledge Management Systems Implementation in Higher Education. Advanced Research in Scientific Areas-International Virtual Conference. Pp. 2078-2081
- [9] Tansley, R.M., Barton, M., Bass, M., Branschovsky, M., McClellan, G., Stuve, D., and Walker, J. (2003). An Open Source Dynamic Digital Repository. *D-Lib Magazine*. 9(1), pp. 1-4.
- [10] Ezema, I.J. (2010), "Building open access institutional repositories for global visibility of Nigerian scholarly publication", *Library Review* Vol. 60 No. 6, 2011 pp. 473-485
- [11] Jain, P. (2010). The Role of Institutional Repository in Digital Scholarly Communications ("unpublished"). Retrieved from [www.library.up.ac.za/digi/docs/jain\\_paper.pdf](http://www.library.up.ac.za/digi/docs/jain_paper.pdf)
- [12] Sarker, F. Davis, H. Tiropanis, T. (2010). The Role of Institutional Repositories in addressing Higher Education Challenges. Pp1-6.
- [13] Shapira, P., Youtie, J., Yogevaran, K., & Jaafar, Z. (2005, May.21). Knowledge economy measurement: Methods, results and insights from the Malaysian knowledge content study. Paper presented at the the Triple Helix 5 Conference on New Indicators for the Knowledge Economy, Turin, Italy.
- [14] Helen Hockx-Yu, (2006) "Digital preservation in the context of institutional repositories", *Program*, Vol. 40 Iss: 3, pp.232 – 243
- [15] Sommerville, I. (2004). Software Requirements. In: Addison-Wesley Software Engineering. 7th ed. London: Pearson Addison Wesley. Chapter 6.
- [16] Mylopoulos, J. (2002). Non-Functional Requirements (or, Quality Factors). Available from <http://www.cs.toronto.edu/~jm/340S/02/PDF2/NFRs.pdf>
- [17] Bawane, N. Srikrishna .V.C. (2010). A Novel Method for Quantitative Assessment of Software Quality. *International Journal of Computer Science and Security*. 3 (6), P.508-512.
- [18] Majid, R.A.; Noor, N.L.M.; Adnan, W.A.W.; Mansor, S., "A survey on user involvement in software Development Life Cycle from
- [19] Levy, D. (2008). Creating highly effective Use Case Diagrams. Available from [http://www.gatherspace.com/static/use\\_case\\_diagram.html#1](http://www.gatherspace.com/static/use_case_diagram.html#1). Last accessed 15th May 2014
- [20] Nath, S. Sridhara, B. Joshi .M. C. Kumar, P. (2008). Intellectual Property Rights: Issues for Creation of Institutional Repository. *Journal of Library and Information Technology*. 28 (5), pp 49-54.

## AUTHORS PROFILE

**Martha Mhongole** is currently a master's student in Information communication science and Engineering at Nelson Mandela African Institution of Science and Technology (NM-AIST) majoring in Information Technology, System Development and Management.

She received a Bachelor degree in Computer Science from Ruaha University College, a constituent college of St. Augustine University of Tanzania in 2011 and she works at same college as a Tutorial Assistant

Her area of interests includes system Analysis, design and development, Database design, Networking, Electronic commerce, computerized Accounting and Knowledge Management.

**Mr. Loserian Saiterie Laizer** graduated in 2000 from the University of Dar es Salaam, Tanzania with BSc. Computer Science majored in Computer

Science and Statistics. He holds an Msc. in IT and Management from Avinashillingam Deemed University of India.

He joined the Nelson Mandela African Institute of Science and Technology in the School of Mathematics, Computational and Communication Science and Engineering in Feb 2011 as an Assistant Lecturer in Information Communication Science and Engineering (ICSE). He holds various certificates offered by local and international bodies.

In 2006 he received a specialized training in Computer Security in Okinawa International Centre - Japan. He has previously worked in various

institutions before joining the Institute. He worked with the former Ministry of Science, Technology and Higher Education as Computer System Analyst for five years. He also served the Bank of Tanzania as a Senior IT Security Administrator for four years.

He has a vast knowledge and experience in IT Security, Business Continuity Management, IT Governance, Statistics, Research Methodology and Finance.



## Finding Untraced Fuzzy Association Rules

F. A. Mazarbhuiya  
College of Computer Science  
Albaha University  
Albaha, KSA

**Abstract**— Fuzzy association rules are rules of the form “If  $X$  is  $A$  then  $Y$  is  $B$ ” where  $X$  and  $Y$  are set of attributes and  $A, B$  are fuzzy sets that describe  $X$  and  $Y$  respectively. In most of fuzzy association rules mining problem fuzziness is specified by users. The users usually specify the fuzziness based on their understanding of the problem as well as the ability to express the fuzziness by natural language. However there exist some fuzziness which cannot be expressed using natural language due its limitation. In this paper we propose a method of extracting fuzzy association rules which cannot be traced by usual methods. We suggest a way of extracting these rules.

**Keywords**- Fuzzy set, Association rules, Fuzzy interval, Certainty factor, Significance factor, Between Operation.

### I. INTRODUCTION

The problem of association rule mining was defined by Agrawal *et al* [4]. Binary association rule mining is to find the relationships between the presences of various items within the baskets. A generalization of the binary association rules is motivated by the fact that a dataset is usually not restricted to binary attributes but also contains attributes with values ranging on ordered scales, such as cardinal or ordinal attributes. Quantitative association rules were defined for dealing with quantitative attributes [5]. In quantitative association rules attribute values are specified by means of subsets, which are typically intervals specified by hard boundaries. This is done by discretizing the domains of quantitative attributes into intervals. Generalizing from hard boundary intervals to soft boundary intervals has given rise to fuzzy association rules. A method for computing fuzzy association rules have been described in [1]. The fuzzy association rules are more understandable to human because of linguistic terms associated with fuzzy sets. The known fuzzy association rules mining techniques may however miss some interesting rules in the process as will be shown here. In this paper, we propose a method, which can extract these missing rules.

The paper is organized as follows. In section II, we discuss briefly about the related works. In section III we review some definitions of basic terms and describe notations and symbols generally used with association rules mining.. In section IV, we discuss the problem that may arise in this method and then describe how to extract the missing rules. Finally in section V, we provide a conclusion and lines for future research.

### II. RELATED WORKS

Replacing crisp sets (intervals) by fuzzy sets (intervals) leads to fuzzy (quantitative) association rules. Thus, a fuzzy association rule is understood as a rule of the form  $A \rightarrow B$ , where  $A$  and  $B$  are now fuzzy subsets rather than crisp subsets of the domains  $D_X$  and  $D_Y$  of two attributes  $X$  and  $Y$  respectively. Each attribute will be associated with several fuzzy sets. In other words, an attribute  $X$  is now replaced by a number of fuzzy attributes rather than by a number of binary attributes. Each element will contribute a vote between 0 and 1 both inclusive to the fuzzy attributes.

The approach made in [1], [2], [6] to generalize the support-confidence measure for fuzzy association rules is to replace set-theoretic operations, namely Cartesian product and cardinality, by corresponding fuzzy set-theoretic operations. In [1] the terms significance and certainty are used instead of support and confidence usually used with non-fuzzy situations.:

### III. TERMS AND NOTATIONS USED

*A. Some basic definitions, terms and notations related to fuzziness*

Let  $E$  be the universe of discourse. A fuzzy set  $A$  in  $E$  is characterized by a membership function  $A(x)$  lying in  $[0, 1]$ .  $A(x)$  for  $x \in E$  represents the grade of membership of  $x$  in  $A$ . Thus a fuzzy set  $A$  is defined as

$$A = \{(x, A(x)), x \in E\}$$

Fuzzy intervals are special fuzzy numbers satisfying the following.

1. there exists an interval  $[a, b] \subset R$  such that  $A(x_0) = 1$  for all  $x_0 \in [a, b]$ , and
2.  $A(x)$  is piecewise continuous.

A fuzzy interval can be thought of as a fuzzy number with a flat region. A fuzzy interval  $A$  is denoted by  $A = [a, b, c, d]$  with  $a < b < c < d$  where  $A(a) = A(d) = 0$  and  $A(x) = 1$  for all  $x \in [b, c]$ .  $A(x)$  for all  $x \in [a, b]$  is known as left reference function and  $A(x)$  for  $x \in [c, d]$  is known as the right reference function. The left reference function is non-decreasing and the right reference function is non-increasing [see *e.g.* [3]].

*B. Some basic definitions related to association rules*

Consider a set  $I = \{i_1, i_2, \dots, i_m\}$  of items, and let a transaction  $t$  (data record) be a subset of  $I$  i.e.  $t \subseteq I$ . Let  $D_X = \{t \in D \mid X \subseteq t\}$  denote the set of transactions in the database  $D$  that contains the items  $X \subseteq I$ . The cardinality of this set i.e.  $|D_X|$  is called the support of  $X$  in  $D$ . Given a minimum threshold  $\sigma$ ,  $X$  is said to be frequent if  $|D_X| \geq \sigma$ . An association rule is a rule of the form  $A \rightarrow B$  where  $A, B \subseteq I$  and  $|D_{A \cup B}| / |D_A| \geq \rho$  where  $\rho$  is another used defined threshold. The support of an association rule  $A \rightarrow B$  is  $|D_{A \cup B}|$ . Sometimes the support is calculated as a fraction of the size of the dataset under consideration. In that case we have  $\text{supp}(A \rightarrow B) = |D_{A \cup B}| / |D|$ . The confidence is the proportion of correct applications of the rule:

$$\text{conf}(A \rightarrow B) = |D_{A \cup B}| / |D_A|$$

Rather than looking at a transaction  $t$  as a subset of items, it can also be seen as a sequence  $(x_1, x_2, \dots, x_m)$  of values of binary variables  $X$ , with domain  $D_X = \{0, 1\}$ , where  $x_j = 1$  if the  $j$ th item,  $i_j$ , is contained in  $t$ , otherwise  $x_j = 0$ .

The association rule mining problem has been extended to handle relational tables rather than transactions of items. In this case the problem is transformed into binary one in the usual way. However a database may contain quantitative attributes (such as age, salary) and in such cases transforming it into binary one will not be possible due to the large size of the underlying domain *e.g.* integers. The discrete interval method [5] divides the quantitative attribute domain into discrete intervals. Each element will contribute support to its own interval. In fact, each interval  $A = [x_1, x_2]$  does again define a binary attribute  $X_A(x)$  defined by  $X_A(x) = 1$  if  $x \in A$  and 0 otherwise. In other words, each quantitative attribute  $X$  is replaced by  $k$  binary attributes  $X_{A_i}$  such that  $X \subseteq \bigcup_{i=1}^k A_i$ .

*C. Significance factor*

The significance factor is calculated by first summing up all votes of each record with respect to the specified item set then dividing it by the total number of records. Let  $A$  be a set of fuzzy sets defined on a set of attributes  $X$ , then the significance factor of the pair  $\langle X, A \rangle$  is calculated as

$$\text{Significance} = \frac{\sum_{t_i \in D} \prod_{x_j \in X} \{\alpha_{a_j}(t_i[x_j])\}}{|D|}$$

where  $t_i$  is the  $i$ -th transaction and  $t_i[x_j]$  gives the value of the  $j$ th attribute in  $t_i$ , and  $m_{a_j}$  is the membership function of  $x_j$ .

$$\alpha_{a_j}(t_i[x_j]) = \begin{cases} m_{a_j \in A}(t_i[x_j]), & m_{a_j} \geq \omega \\ 0 & \text{otherwise} \end{cases}$$

where  $\omega$  is a user specified threshold.

#### D. Significance of a fuzzy association rule

The significance of a fuzzy association rule  $A \rightarrow B$ , where  $A$  and  $B$  are fuzzy subsets is defined in [1] as the ratio of the sum of the memberships  $A(x)$  and  $B(y)$ , provided  $(x, y) \in D$ , to the total number of transactions in  $D$ . i.e.

$$\text{significance}(A \rightarrow B) = \frac{\sum_{(x,y) \in D} A(x) \otimes B(y)}{|D|}$$

where  $\otimes$  is  $\Pi$  i.e. the *mul* operator.

#### E. Certainty of a fuzzy association rule

The certainty of a fuzzy association rules  $A \rightarrow B$ , is defined as

$$\text{certainty}(A \rightarrow B) = \frac{\sum_{(x,y) \in D} A(x) \otimes B(y)}{\sum_{(x,y) \in D} A(x)}$$

#### F. Definition

A fuzzy association,  $A \rightarrow B$  is said to hold if  $\text{certainty}(A \rightarrow B)$  is greater than or equal to  $\text{min\_cert}$  and  $\text{significance}(A \rightarrow B)$  is greater than or equal to  $\text{min\_sig}$  where the thresholds  $\text{min\_cert}$  and  $\text{min\_sig}$  are provided by the user.

### IV. LIMITATION OF EXISTING METHOD AND NEW APPROACH

In this section we show how some interesting fuzzy association rules may be missed out due to the way in which a user has specified the input fuzzy sets. For example let us consider fuzzy sets defined on an attribute specifying the time in hours at which an event has occurred. The fuzzy sets might be mid-night, morning, afternoon, evening and night as described by the following diagram.

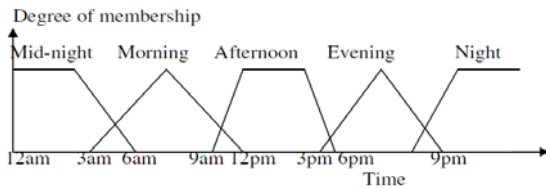


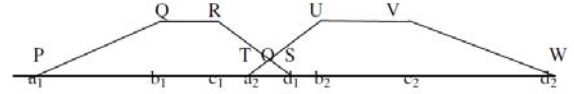
Fig. Definition of linguistic terms for attributes *Time-of-occurrence*.

Consider two consecutive fuzzy intervals mid-night and morning. The points that contribute to both the intervals i.e the transaction between 3 a.m and 6 a.m will contribute much less than 1 to both the intervals. So if there are a reasonable number of events taking place in this period then due to the manner in which their supports are calculated, their contributions will be less in both the intervals. Thus some of the rules may be left undiscovered. This may happen in the case of non-fuzzy intervals also. For example if all the transactions are uniformly distributed in the interval  $[1, 4]$  and if the user-specified intervals are  $[0, 2]$  and  $[2, 4]$ , then it is not possible to identify  $[1, 4]$  as a frequent interval although it might be frequent if we would have considered  $[1, 4]$  as an interval. If we consider the fuzzy interval, which is in between Midnight and Morning (the between operation is defined in section-4.1), then the fuzzy interval lying between Midnight and Morning may turn out to be frequent.

The known methods are very much user dependent due to fact that the fuzzy intervals are supplied by the domain expert. The domain expert may not have sufficient knowledge about the datasets. So he will supply the fuzzy intervals according to his limited understanding of the dataset and fuzzy intervals, which can be expressed using linguistic terms. So there is an every possibility that some of the association rules may be left undiscovered. Actually the time stamps lying between the fuzzy intervals are given less emphasis, which may not turn out to be appropriate always.

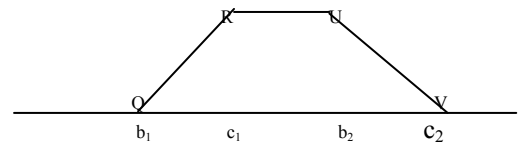
#### A. Between operation

Let  $A$  and  $B$  are two adjacent fuzzy intervals specified by user. They are called adjacent if  $A(x) \cap B(x) \neq \emptyset$  for some  $x \in E$  and there is no user specified fuzzy interval in between. We define an operation called “Between”, which takes  $A$  and  $B$  as input and returns a fuzzy interval covering the portion between  $A$  and  $B$  with membership value less than 1. To illustrate this we take  $A = [a_1, b_1, c_1, d_1]$  and  $B = [a_2, b_2, c_2, d_2]$  be two adjacent fuzzy intervals and  $A(x) \cap B(x) \neq \emptyset$  for some  $x \in E$ , they are shown in the figure below:



Here ROU is the portion neither belonging  $A$  with full membership nor belonging to  $B$  with full membership. Our between operation takes the above two intervals as input and returns a fuzzy interval with core RU. We denote it by  $A(B)B = [b_1, c_1, b_2, c_2]$

$$\text{where } (A(B)B)(x) = \begin{cases} (x-b_1)/(c_1-b_1), & b_1 \leq x \leq c_1 \\ 1, & c_1 \leq x \leq b_2 \\ (c_2-x)/(c_2-b_2), & b_2 \leq x \leq c_2 \end{cases}$$



Given an underlying data set and fuzzy sets as input, our method is to consider the fuzzy set formed by joining consecutive fuzzy intervals as discussed above together with the input fuzzy intervals while calculating the significance factor and finally finding the association rules. In this process each and every time-stamp in the whole duration under consideration is given equal importance and every such time stamp contributes 1 to at least one fuzzy interval under consideration. In this process obviously all existing association in the data set will be detected.

#### A. SUMMARY AND LINES FOR FUTURE WORKS

An approach to finding fuzzy association rules that may be missed by other existing methods are discussed here. The algorithm extracts all the fuzzy association rules extracted by other methods and possibly some more. The other methods do not consider the region lying between two consecutive fuzzy intervals specified by user. If a record falls in this region its contribution will be less in both the fuzzy intervals. If sufficient number of records is lying in this region, the sum of their contributions may still be less than the user specified threshold and hence some association rules associated with such regions will be left undiscovered. In this paper, we proposed an approach, which takes into consideration all such regions falling between two consecutive fuzzy intervals specified by the user. So obviously this method gives all the fuzzy association rules specified by user plus it also extracts some extra association rules which is the beyond the scope of user. Future works may be done in the following two lines

- attempt to find rules for fuzzy data
- instead of taking the fuzzy intervals as input attempts may be made to extract the fuzzy intervals from the dataset in a natural way.

#### REFERENCES

- [1] C. Man Kuok, A. Fu, and M. Hon Wong. Mining Fuzzy Association Rules in databases, SIGMOD Record, 27:41-46, 1998.
- [2] H. Prade, E. Hullermeir and D. Dubois, A Note on Quality Measures for Fuzzy Association Rules, In Proceedings IFSA-03, 10<sup>th</sup> International Fuzzy Systems Association World Congress. LNAI 2715, Istanbul (July 2003) 677- 684.

- [3] J. Klir and B. Yuan; Fuzzy Sets and Logic Theory and Application, Prentice Hill Pvt. Ltd.(2002)
- [4] R. Agrawal and R. Srikant. Fast algorithms for mining association rules, In Proceedings of the 20<sup>th</sup> conference on VLDB, Santiago, Chile, 1994.
- [5] R. Srikant and R. Agrawal, Mining quantitative association rules in large relational tables, 1995.
- [6] Wai-Ho Au and Keith C. C. Chan; Mining Fuzzy Association Rules, Proc. of the Sixth Int'l Conf. on Information and Knowledge management, 1997, 209-215.

University, India. After this he obtained the Ph.D. degree in Computer Science from Gauhati University, India. He had been working as an Assistant Professor in College of Computer Science, King Khalid University, Abha, Kingdom of Saudi Arabia from October 2008 to September 2011. Currently he is serving as Assistant Professor at College of Computer Science and IT, Albaha University, Kingdom of Saudi Arabia. His research interest includes Data Mining, Information security, Fuzzy Mathematics and Fuzzy logic.

#### AUTHORS PROFILE



**Fokrul Alom Mazarbhuiya** received B.Sc. degree in Mathematics from Assam University, India and M.Sc. degree in Mathematics from Aligarh Muslim

# A Novel Approach to Address Information Leakage Attacks Based on Machine Virtualization

Omar Hussein<sup>1</sup>, Nermin Hamza<sup>2</sup>, Hesham Hefny<sup>3</sup>

*Computer and Information Sciences Department  
Institute of Statistical Studies and Research, Cairo University, Egypt*

<sup>1</sup>ohusseins@gmail.com

<sup>2</sup>nermin.hamza@cu.edu.eg

<sup>3</sup>hehefny@ieee.org

**Abstract**—In a traditional non-virtualized computer system the whole software stack is highly vulnerable to security breaches. This is mainly caused by the coexistence of deployed security systems in the same space as the potentially compromised operating system and applications that often run with administrative privileges. In such a structure, compromising, bypassing, disabling, or even subverting deployed security systems become trivial. Machine virtualization provides a powerful abstraction for addressing information security issues. Its isolation, encapsulation, and partitioning properties can be leveraged to reduce computer systems' susceptibility to security breaches. This paper demonstrates that machine virtualization when employed and synthesized with cryptography would preserve information confidentiality even in an untrusted machine. It presents a novel information security approach called Virtualized Anti-Information Leakage (VAIL). Its objective is to thwart malicious software and insiders' information leakage attacks on sensitive files after decryption in potentially compromised computer systems. VAIL's defenses are evaluated against a variety of information leakage attacks including: (1) direct attacks launched on sensitive files from an untrusted virtual machine, and a compromised virtual machine monitor; and (2) indirect attacks exploiting covert storage and timing channels. Based on the security evaluation, it is concluded that VAIL effectively complied with the security requirements, and met its objective.

**Index Terms**—Information Security; Information Leakage; Machine Virtualization; Malicious Software; Insider Threat

## I. INTRODUCTION

Information leakage attacks represent a serious threat for their widespread and devastating effects. Significance of such attacks stems from the fact that they are committed by an organization's authorized computer users, and/or processes executing on their behalf. The diverse avenues that could be exploited to carry out these attacks add another barrier towards addressing them.

In this paper focus is driven towards malicious software (malware) and the insider threat for being the most prominent perpetrators of information leakage attacks. Malware continues to form a major threat, whilst the insider threat is prevailing and represents a challenging unsolved problem for two main reasons: (1) insiders possess deep understanding of the targeted vulnerable processes; and (2) they are aware of systems'

unpatched security vulnerabilities. Consequently, addressing malware and the insider threats is a key security requirement.

To highlight the problem area, the following example is presented. An accountant created a spreadsheet file to maintain the company's bank account number, balance, total credits and debits, etc. He/she regularly downloads renewal statements and statements of account from the bank's website and edits the spreadsheet file. To prevent unauthorized disclosure and propagation of such sensitive financial information, the company mandates, according to its security policy, encrypting sensitive files. However, after decryption, sensitive files are still exposed to information leakage attacks. New undetected malware may attempt to leak out the file's contents after capturing its decryption password and/or opening it. In addition, being an authorized user, the accountant, or any of his co-workers may exploit their privileges to leak out such sensitive information to the company's competitors for personal or financial gain.

This paper presents a novel information security approach called Virtualized Anti-Information Leakage (VAIL). Its objective is to thwart malware and insiders' information leakage attacks on sensitive files after decryption in potentially compromised computer systems. VAIL's basic idea lay in the method machine virtualization and cryptography are synthesized and employed to achieve this objective. VAIL's defenses are evaluated against a variety of information leakage attacks including: (1) direct attacks launched on sensitive files from an untrusted virtual machine, and a compromised virtual machine monitor; and (2) indirect attacks exploiting covert storage and timing channels.

The remainder of this paper is organized as follows: Section II briefly explains machine virtualization and its security-related advantages. Section III provides an overview of the previous work that exploited machine virtualization in information security applications. Section IV presents VAIL; the security requirements, threat model and assumptions, evaluation of design alternatives, VAIL structure and overview of its components, its encryption scheme, and operation. Section V evaluates VAIL's defenses against direct and indirect information leakage attacks. Finally, Section VI concludes the paper.

## II. MACHINE VIRTUALIZATION

Machine virtualization is a computer system abstraction that aims at detaching workloads (i.e., operating systems (OSs) and applications) and data from the functional side of the physical hardware [29]. Through machine virtualization, multiple isolated *guests* (called virtual machines (VMs)) having heterogeneous unmodified OSs concurrently run on top of a virtual machine monitor (VMM), which resides directly above the *host* hardware (Figure 1). A VM behaves like a separate computer, in which its virtual resources are a subset of the machine's physical resources.

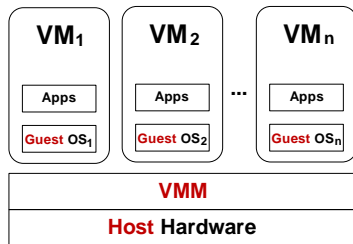


Fig. 1. Machine virtualization structure

The VMM manages, mediates access to, and creates multiple instances of a physical resource than exists in reality. It presents to each guest a picture of the resource that corresponds to its own context. It provides the necessary mapping between the physical resources and VMs' virtual devices. It intercepts guests' privileged instructions on virtual devices and handles them before they are executed by the physical hardware under its control. Being in the highest privilege level, the VMM is able to isolate itself from the VMs and isolate the VMs from each other.

### A. Security-Related Advantages

Isolation, encapsulation, minimal code size, and partitioning properties of machine virtualization can be leveraged to reduce computer systems' susceptibility to security breaches as clarified below.

#### - Minimized Attack Surface through Isolation

Machine virtualization provides strong isolation between the VMM and its VMs, and between the VMs. This minimizes the attack surface to be confined to the potentially compromised VM(s), and prevents adversaries from expanding their attacks to adjacent VMs sharing the same host hardware. Furthermore, the isolation property could be exploited to prevent adversaries from disabling or even subverting the desired security functionality. Security applications could be deployed in dedicated VM(s) having privileged access to other potentially compromised VMs and to the physical hardware resources through the VMM.

#### - VM Secure State Restoration through Encapsulation

The VMM can encapsulate the execution state of a VM and resume the execution of a pre-configured VM image. Through leveraging the encapsulation property, security administrators could instantly restore a secure VM image by rolling a compromised VM back to some previously checkpointed secure state; then resume the VM normally. Such capability facilitates systems administration and simplifies the lengthy complicated setup procedure needed for a physical server.

#### - Maximized Security through Thin VMMs

OSs are very complex and large. Their sizes fall in the range of tens of millions of lines of code (LOC), which makes it inappropriate to consider them secure [3]. Regarding VMMs, Xen [31], for instance, is implemented in under 50,000 LOC, whereas seL4 microkernel [18] has 8,700 lines of C code and 600 lines of assembler. Other examples include BitVisor [28] that comes at 21,582 LOC, and SecVisor [27] that comes at 1,739 and 1,112 LOC for each of its two versions. Number of security vulnerabilities is directly proportional to the code size. Vulnerability reports confirmed this fact; showing, for instance, only 72 security vulnerabilities for Xen 4.x [25]; whereas showing 310 for Microsoft Windows 7 [26].

#### - Enhanced VM Isolation through Static Partitioning of Resources

The VMM has the property of partitioning the system resources among its VMs. Partitioning could be accomplished statically or dynamically. In dynamic partitioning, resources are allocated to and de-allocated from VMs as needed during their execution. In static partitioning, each VM is provided access only to its fixed allocated resources, thereby providing strengthened isolation between adjacent VMs sharing the same host hardware [29].

## III. RELATED WORK

Information security researchers presented a number of applications that focused mainly on leveraging the isolation and encapsulation properties of machine virtualization. These applications could be categorized into: (1) malware analysis, detection, and prevention; (2) intrusion analysis and detection; and (3) digital forensics as follows.

### A. Malware Analysis, Detection, and Prevention

MAVMM [23] is a lightweight and special-purpose VMM for malware analysis. MAVMM supports a single guest OS and makes use of *hardware-assisted virtualization* to minimize VMM's code base and make it harder for malware to detect VMM existence. It has the capability of extracting many features from programs running inside the guest OS, such as fine grained execution trace, memory pages, system calls, disk and network accesses.

VMWatcher [16] uses VM introspection (VMI) to detect rootkits. It captures the current state of a guest OS and compares it with the state reported by the guest OS itself to detect any inconsistencies in the data structure of currently executing programs. Patagonix [21] relies on a monitor implemented in the VMM, and a separate VM in order to detect binaries covertly executing in a VM. As input, Patagonix requires information about the binaries it will identify that is represented in a list of *known* binaries. Any other binaries not in the list are identified as unknown. The results of the processes' identity-checking are sent to the user, who can compare Patagonix's report on currently executing binaries with those reported by the guest OS.

SecVisor [27] and NICKLE [24] are VMM-based systems that aim at preventing rootkits from execution. SecVisor is a tiny special-purpose VMM that supports only a single central processing unit (CPU) core and a single VM running a commodity OS. It leverages hardware-assisted virtualization to write-protect kernel code pages in the guest OS's memory, and approves loading of kernel modules only in case they have been pre-incorporated in a *whitelist* of computed cryptographic hashes. NICKLE is a trusted VMM; it maintains a separate guest-inaccessible *shadow memory* for a running VM to store authenticated kernel code. At runtime, NICKLE transparently redirects guest kernel instruction fetches to the shadow memory. As a result, only authenticated kernel code will be fetched for execution, and rootkits that normally require executing their own attack code will not be executed.

### B. Intrusion Analysis and Detection

Intrusions could be analyzed by logging and replaying VM execution. In case a VM was subject to an intrusion, *logging records* could be used to analyze the break-in process by replaying the VM instructions. This could help in determining the attack source, cause, and effects. ReVirt [5] is a VMM-based system that adopts such an approach. As an advancement to ReVirt, SMP-ReVirt [6] brings ReVirt's logging and replaying functionality to multiprocessor VMs running on commodity hardware.

Livewire [10] is a VMI-based intrusion detection system (IDS). It makes use of two VMs running above a VMM; an IDS VM, and the monitored VM. Through VMI, the VMM enables the IDS VM to inspect the state of the monitored VM, and monitor the interactions between its guest OS and virtual hardware. Livewire accomplishes its intended functionality through three components located in the IDS VM: (1) the OS interface library that interprets the hardware state exported by the VMM in order to provide an OS-level view of the monitored VM; (2) the policy engine that consists of a framework for building policies needed for the IDS; and (3) policy modules that implement these policies.

IntroVirt [17] is an IDS. It leverages *vulnerability-specific predicates* to detect whether adversaries had exploited specific security vulnerabilities before releasing the relevant patches, or in the period between the patches release time and the

application time. These predicates are written by software patch author(s) and executed at a particular invocation point within the OS or application code. IntroVirt uses ReVirt [5] to replay execution of a monitored VM. During replay, IntroVirt monitors the security predicates to find out if any of the announced vulnerabilities have already been exploited. It relies on VMI to monitor predicates checks, and inspect the state of the monitored VM. When a breakpoint is encountered, the VMM checkpoints the VM, invokes the predicate code and then rollbacks the VM back and resumes execution. Rolling back the VM ensures that executing a predicate does not affect the VM.

### C. Digital Forensics

VMI allows digital forensics investigators to carry out live VM analysis, where the dynamic state of a target VM can be obtained without allowing it to detect that it is being monitored. For instance, the virtual introspection for Xen suite of tools [13] contains a set of utilities built over an inspection library that can be used from a Xen administrative domain to examine a running VM. Such capability can be used to reveal relevant forensics data, or discrepancies between the guest OS and its view from the VMM perspective that could be caused by malware such as rootkits.

## IV. THE PROPOSED INFORMATION SECURITY APPROACH

This section presents VAIL; a novel information security approach that aims at thwarting malware and insiders' information leakage attacks on sensitive files after decryption in potentially compromised computer systems. It begins by enumerating the intended security requirements that VAIL should meet. It proceeds by describing the threat model, evaluating the design alternatives, and justifying choices made. It explains VAIL structure and overviews its components. It explains its encryption scheme, and, finally, illustrates its operation.

### A. Security Requirements

VAIL should meet five security requirements:

- 1) *Provide an OS-independent trusted boundary.* Remove the OS from users trust base, whilst preserving information confidentiality after file decryption in a potentially compromised computer system.
- 2) *Prevent circumvention.* Continue to function as intended in untrusted computer systems, and provide protection against *spoofing attacks*.
- 3) *Achieve 256-bit security strength.* A *brute-force attack* on a cryptographic key would require  $2^{256}$  steps. A step refers to performing a single encryption operation on a given plaintext value with a key, then comparing the result with a given ciphertext value [2].



- 4) *Provide simplicity in usage.* Provide automatic and transparent cryptographic operations to users, and limit their interactions with the proposed approach to increase its simplicity and usability.
- 5) *Provide compatibility with existing OSs and hardware.* Provide direct applicability to commercial off-the-shelf OSs and applications, and without any special or additional required hardware.

#### B. Threat Model and Assumptions

It is assumed that OSs and applications are untrusted and could be compromised. They will always contain exploitable security vulnerabilities. OS-authenticated users are deemed untrusted as well. They do not have full access to any trusted components (explained in Subsection D). Malware and insiders are the main adversaries. At runtime, they may provide spoofed user interfaces (UIs), and attempt to spy on users' authentication information. They may attempt to leak out information through: (1) breaching the isolation imposed on the encrypted sensitive files while in storage; (2) brute-forcing the encryption keys after compromising the VMM; or (3) exploiting covert storage and timing channels.

#### C. Evaluation of Design Alternatives

This subsection briefly evaluates VAIL's design alternatives; it justifies choice of: (1) virtualization type; (2) CPU virtualization technique; (3) VMM design; (4) stored data encryption approach; and (5) cryptographic algorithm and key length.

##### - Choice of Virtualization Type

In order to securely function in a potentially compromised computer system, isolation from adversaries is critical. Machine virtualization contributes in achieving this requirement by leveraging its isolation security-related advantage. According to the threat model, the OS is deemed untrusted, therefore *OS virtualization* is discarded. *Application virtualization* is discarded as well since each sandboxed application runs above the VMM, which runs as an application on top of the (untrusted) OS. Consequently, VAIL will adopt machine virtualization.

##### - Choice of CPU Virtualization Technique

CPU virtualization techniques involve how privileged and user instructions are handled. They are classified into *full virtualization*, *paravirtualization*, and *hardware-assisted virtualization* [11]. Full virtualization virtualizes the host hardware to enable concurrent execution of multiple heterogeneous unmodified guest OSs using *interpretation* or *dynamic binary translation*; however, causing performance degradation. Eradicating the emulator's code in paravirtualization yields a virtualized system with a smaller footprint size and reduced overhead. However, it requires deep error-prone manual modification to guest OSs [14]. *Pre-virtualization* is an automated form of paravirtualization in which unmodified guest OSs run

above an intermediary that, in turn, runs above the VMM [20]. Recently, Intel processors [15] provided support for full virtualization in hardware, thereby eliminating the need for interpretation and dynamic binary translation. According to the security requirements, VAIL should be directly applicable to commercial off-the-shelf OSs and applications, without any special or additional required hardware. Consequently, since pre-virtualization brings the benefits of both full virtualization and paravirtualization, it is the chosen CPU virtualization technique.

##### - Choice of VMM Design

A *microkernel* is chosen over a *monolithic* VMM for the following reasons [30]: (1) it offers a minimal layer over the hardware platform by focusing on providing basic system services, and eliminating OS kernel nonessential modules (e.g., device drivers); (2) it is safer and more reliable, since most services run as user rather than kernel processes; and (3) it is easier to validate, maintain, extend, and port from one hardware design to another. Consequently, taking account of the preceding considerations, VAIL will adopt machine virtualization in which VMs running unmodified OSs run above a pre-virtualization layer installed on top of a microkernel that runs directly above the host hardware (VAIL structure is explained in Subsection D).

##### - Choice of Stored Data Encryption Approach

Encryption approaches to protect data at rest fall into three categories: (1) whole disk encryption; (2) file system-level encryption; and (3) file-level encryption as follows:

###### • Whole Disk Encryption

In whole disk encryption (WDE), such as Microsoft's BitLocker Drive Encryption [22], the cryptographic solution is placed at the hardware device itself, or in the appropriate device driver. Entire disks or partitions including all system and user files are encrypted using user-supplied passwords. WDE suffers from several drawbacks: (1) insiders could leak out information during unattended sessions, or through stolen decryption keys; (2) WDE is vulnerable to *cold boot attacks*, where cryptographic keys could be stolen from the dynamic random access memory by exploiting its ability to retain its contents for several seconds after machine shutdown [12]; and (3) the computer system incurs the overhead of encrypting and decrypting contents of entire disks including unclassified information.

###### • File System-Level Encryption

In file system-level encryption, cryptography is performed at the file system layer of the OS. However, its implementations suffer from inconsistency of encryption algorithms across successive file systems. For example, files encrypted by the encrypting file system (EFS) of Windows XP OS Service Pack 1 or later cannot be decrypted using EFS



incorporated in earlier Windows versions. This is because recent EFSs employ the Advanced Encryption Standard (AES) cryptographic algorithm [8], whereas older EFSs support only the expanded Data Encryption Standard and the Triple Data Encryption Standard algorithms [7].

- File-Level Encryption

In file-level encryption (FLE) chosen files are manually encrypted and decrypted with a specified key (e.g., UNIX *crypt* utility). Despite its simplicity, it has its own cons [19]: (1) there is no means to ensure users' commitment to delete plaintext versions of sensitive files after encryption, and re-encrypt them after each decryption; (2) after decryption, files are exposed to the risk of malicious manipulation; and (3) as the number of sensitive files increases, the overhead to manually decrypt and re-encrypt these files increases as well.

However, VAIL will adopt the FLE approach for the following reasons: (1) FLE prevents adversaries from compromising other still encrypted files in case a file's key was revealed or stolen; (2) VAIL intends to provide automatic and transparent cryptographic operations to its users, where sensitive files are never stored in plaintext, they are automatically encrypted at creation time, and forcibly re-encrypted at close time; and (3) an organization's sensitive files are naturally expected to be clearly identified according to its security policy, much fewer compared to total number of business files, and with low update frequency.

- Choice of Cryptographic Algorithm and Key Length

VAIL will adopt the AES symmetric-key block cipher for three main reasons: (1) it is an approved Federal Information Processing Standard (FIPS) cryptographic algorithm for the protection of sensitive information [8]; (2) it is more efficient than the other FIPS approved cryptographic algorithm (i.e., the Triple Data Encryption Algorithm) [2]; and (3) it is an industry-leading encryption algorithm used in several versions of Windows OS to encrypt entire OS drives [22]. A key length of 256 bits is chosen for being necessary to achieve the required security strength.

#### D. VAIL Structure and Overview

In order to provide the intended protection, some trusted components are necessary. Figure 2 illustrates VAIL structure in which trusted components are drawn in solid lines, whereas untrusted ones are drawn in dashed lines. VAIL operates using two VMs; the *Vulnerable VM*, and the *Quarantined VM*. Both VMs run concurrently above the VMM (a microkernel) through a pre-virtualization layer. The VMM resides directly above the host hardware. The Vulnerable VM runs the vulnerable OS. It represents the user's personal computer; it may be compromised and is untrusted. The Quarantined VM is responsible for storing sensitive files. It has no user applications installed on it, and is trusted. The VMM, the pre-virtualization layer, and the hardware are considered trustworthy.

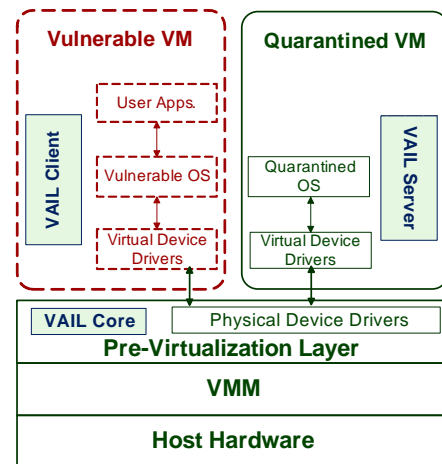


Fig. 2. VAIL structure

Through VAIL a computer system has two states of operation that are named after the VM currently in focus: (1) *vulnerable state* (i.e., the low-confidentiality state); and (2) *quarantined state* (i.e., the safe editing high-confidentiality state). While the computer system is in vulnerable state, the vulnerable OS has access to all devices and can communicate freely over the network. While the computer system is in quarantined state, the vulnerable OS is blocked from sending output through the network or to external storage devices. To provide safe access to sensitive files, VAIL includes three more components to the above structure: *VAIL Core*, *VAIL Server*, and *VAIL Client* as follows:

- 1) *VAIL Core*. It resides in the pre-virtualization layer. It verifies predefined escape sequences that are needed to create sensitive files and manage transition between vulnerable state and quarantined state. It saves a vulnerable state before transition to quarantined state, and restores it back after the user closes a sensitive file.
- 2) *VAIL Server*. It runs inside the Quarantined VM. It creates, encrypts, and stores sensitive files. It decrypts and opens them after authenticating VAIL users by verifying their supplied passwords through *VAIL Password-Based Key Derivation and Verification (VPKDV)* module. VPKDV also derives keys from user-supplied passwords; these keys are used to encrypt files' encryption keys. VAIL Server stores the current vulnerable state before decrypting and opening a sensitive file in quarantined state, and re-encrypts sensitive files upon returning back to vulnerable state.
- 3) *VAIL Client*. It runs in the Vulnerable VM. It is responsible for interacting with the user, storing links to sensitive files, and salting user-supplied passwords. A comprehensive explanation of VAIL operation is provided in Subsection F.

### E. VAIL Encryption Scheme

VAIL employs AES-256 in cipher block chaining (CBC) mode of operation. The following subsections explain key generation, file encryption, and user authentication for file decryption processes.

#### - Key Generation

VAIL employs the AES-CBC for two purposes. Firstly, to encrypt sensitive files using 256-bit keys called VAIL File Cipher Keys (VFCKs). Secondly, to encrypt each VFCK, using a user-supplied password-based 256-bit key called VAIL Master Cipher Key (VMCK). VFCK and VMCK generation as follows:

- 1) Whenever a sensitive file  $i$  is created,  $VFCK_i$  is assigned a 256-bit value generated from invoking a pseudo-random number generator (PRNG) running in the Vulnerable VM.

$$VFCK_i = PRNG()$$

- 2) Whenever  $User_x$  creates a VAIL account, alters or uses his/her password, it is concatenated with a fresh unique 256-bit random bitstring (called the salt) that is generated from invoking  $PRNG()$ . It is then iteratively hashed through VPKDV to generate the  $VMCK_x$ . VPKDV is a module located in the VAIL Server to: (1) derive keys from user-supplied passwords (Figures 3 and 4), these keys are used to encrypt VFCKs; and (2) authenticate users by verifying their supplied passwords (Figures 6 and 7).

$$VMCK_x = CrtModPsw(u, p, z)$$

As illustrated in Figure 3,  $User_x$  creates a password  $P_x$  of eight 7-bit ASCII characters (i.e., 56 bits). It is salted with  $S_x$  that is assigned a 256-bit output from  $PRNG()$ . The salted password is then iteratively hashed with the secure hash algorithm SHA-256 [9] to compute a salted hash that is assigned to  $VMCK_x$ . The user identifier (UID), the salt, and the salted hash are then stored in the password file (PSF) that resides in the Quarantined VM. VPKDV functions to salt and iteratively hash user-supplied passwords are defined in Figure 4.

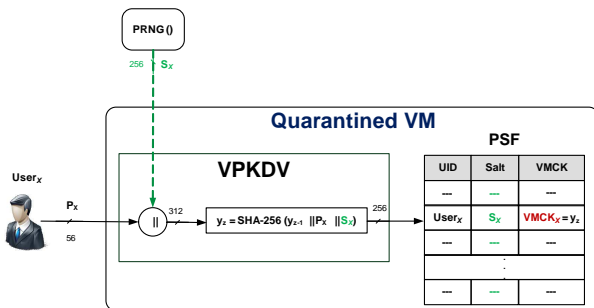


Fig. 3. The process of deriving VMCK using VPKDV

```

function CrtModPsw( $u, p, z$ ) // Create or modify the password
input:  $u$  UID
input:  $p$  password
input:  $z$  number of iterations
output:  $x$  VMCK
 $s = PRNG()$  // salt from a PRNG function
 $x = SaltIterHashPsw(s, p, z)$  // call password salting and iterative hashing function
write (PSF,  $u, s, x$ ) // insert into the password file
return  $x$ 

function SaltIterHashPsw( $s, p, z$ )
input:  $s$  salt
input:  $p$  password
input:  $z$  number of iterations
output:  $x$ 
 $y_0 \leftarrow \epsilon$  // Begin with the empty string in the 1st iteration
for  $a = 1, \dots, z$  do
     $y_a \leftarrow SHA-256(y_{a-1} || s || p)$  // iteratively salt and hash the password 'z' times
end do
 $x \leftarrow y_z$ 
return  $x$  // salted and iteratively hashed password
    
```

Fig. 4. VPKDV functions to salt and iteratively hash user-supplied passwords

#### - File Encryption

A high-level view of VAIL file and key encryption process is as follows (Figure 5):

- 1) The plaintext of file  $i$  ( $PF_i$ ) is encrypted using AES-CBC and the 256-bit  $VFCK_i$  as the key

$$CF_i = AES-CBC_{VFCK_i}(PF_i)$$

- 2)  $VFCK_i$  is encrypted using AES-CBC and the 256-bit  $VMCK_x$  as the key. The 128 most significant (MS) bits and the 128 least significant (LS) bits of the encrypted  $VFCK_i$  are then inserted in FCF\_a and FCF\_b respectively. These two fields are located in the header of the ciphertext of file  $i$  ( $CF_i$ ).

$$write(FCF\_a, Extract\_MS(AES-CBC_{VMCK_x}(VFCK_i)))$$

$$write(FCF\_b, Extract\_LS(AES-CBC_{VMCK_x}(VFCK_i)))$$

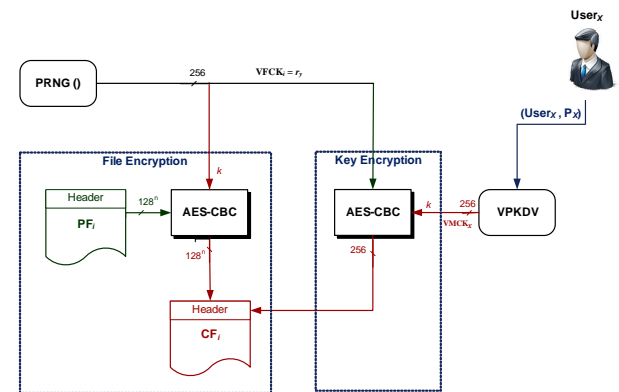


Fig. 5. High-level view of VAIL file and key encryption process

### - User Authentication for File Decryption

Files are decrypted only to users who successfully authenticate themselves to VAIL through VPKDV module. As depicted in Figure 6, when User<sub>x</sub> requests to open a sensitive file, he/she provides a unique UID and a password (P'<sub>x</sub>). VPKDV uses the UID to index into the PSF and retrieve the user's S<sub>x</sub>. It verifies P'<sub>x</sub> by computing VMCK'<sub>x</sub>. It concatenates P'<sub>x</sub> with S<sub>x</sub>, iteratively hashes the result, and compares it with the correspondent previously stored VMCK<sub>x</sub>. If VMCK'<sub>x</sub> and VMCK<sub>x</sub> are identical, then the file decryption process begins. Otherwise, VPKDV terminates the decryption process. The VPKDV function to authenticate users through verifying their claimed passwords is defined in Figure 7.

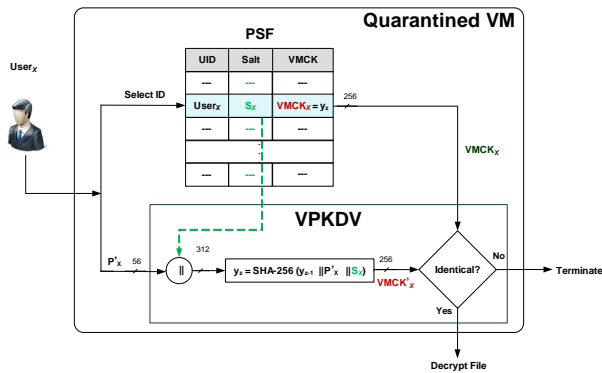


Fig. 6. Authenticating a user

```
function UsePsw(u', p', z)           //On Usage of a password
input: u' claimant UID
input: p' claimant password
input: z number of iterations
a ← seek (PSF, "u == u' ")           //use the UID to index into the password file
(s, x) ← PSF[a]                     //retrieve the user's stored salt and VMCK
x' = SaltIterHashPsw (s, p', z)      //compute the claimant's VMCK
if x == x' then
    begin the file decryption process
Else
    Terminate
end if
return true
```

Fig. 7. VPKDV function to authenticate users by verifying their passwords

### F. VAIL Operation

This subsection details VAIL's operation to clarify the interactions between VAIL Core, VAIL Server, and VAIL Client to thwart malware and insiders' information leakage attacks on sensitive files after decryption. Some steps involve requesting the user to enter escape sequences as a security measure to prevent spoofing by the untrusted Vulnerable VM (see Subsection A in Section V).

### - Creating a Sensitive File

This process aims at assuring that sensitive files are always stored encrypted in the Quarantined VM. Upon receiving a user request from the Vulnerable VM to create a sensitive file, the following steps take place (Figure 8):

- 1) VAIL Core disables the Vulnerable VM's network, and external storage devices drivers (e.g., compact disk-rewriteable), and starts the Quarantined VM's Dynamic Host Configuration Protocol (DHCP) client service.
- 2) VAIL Client captures the file name from the user.
- 3) VAIL Client sends a request to VAIL Server to create a new sensitive file.
- 4) VAIL Server checks sufficiency of free space on the Quarantined VM, if it is adequate, it approves the request.
- 5) VAIL Client requests the user to enter a predefined escape sequence (e.g., Alt + Del) that will be captured by VAIL Core.

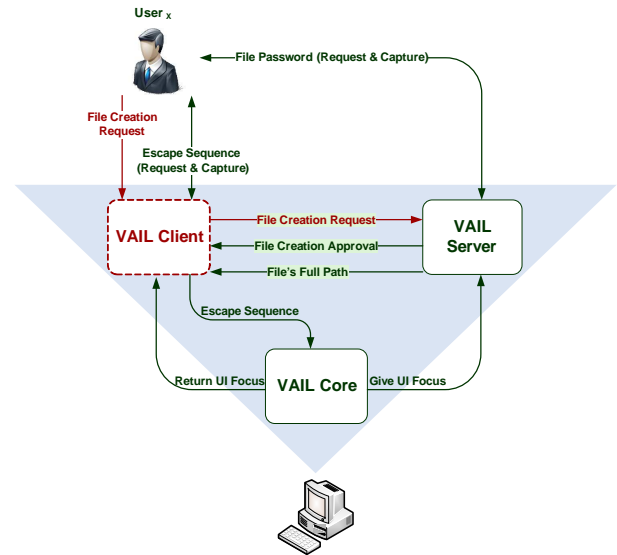


Fig. 8. VAIL process of creating a sensitive file

- 6) Upon receiving and verifying the escape sequence, VAIL Core gives the UI focus to VAIL Server.
  - If VAIL Core does not receive the escape sequence, or receive an incorrect one from VAIL Client, then it will not give the UI focus to VAIL Server. Instead, it will display an alert noting that it is an unsafe process, terminates it, and undo the first four steps.
- 7) VAIL Server creates the file in the Quarantined VM, encrypts it, and requests the user to select a password that will be salted and iteratively hashed to encrypt the file's key (i.e., VFCK).
- 8) VAIL Server sends to VAIL Client a link to the file,

which is its full path.

- 9) VAIL Server requests VAIL Core to return the UI focus back to VAIL Client.
- 10) VAIL Core stops the DHCP client service on the Quarantined VM, re-enables Vulnerable VM's network and external storage devices drivers, and returns the UI focus back to VAIL Client.

#### - Opening a Sensitive File in Quarantined State

This process aims at preventing adversaries from: (1) capturing users' files' decryption passwords; and (2) leaking out sensitive information externally after opening sensitive files. At this stage, one or more sensitive files reside in the Quarantined VM, and have links to them in VAIL Client. The transition process from vulnerable state to quarantined state comprises the following steps (Figure 9):

- 1) VAIL Core disables Vulnerable VM's network and external storage devices drivers, and starts Quarantined VM's DHCP client service.
- 2) VAIL Client sends VAIL Server a user request to open a sensitive file. The request includes the file's full path.
- 3) Upon VAIL Server's approval, VAIL Client requests the user to enter a predefined escape sequence that will be captured by VAIL Core to perform the state transition.

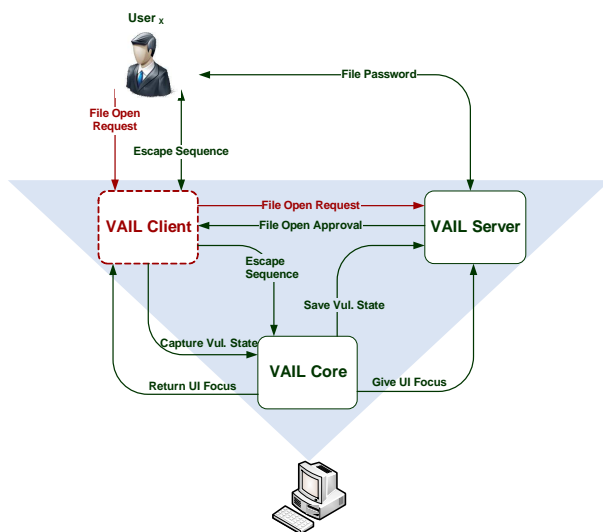


Fig. 9. VAIL process of opening a sensitive file

- 4) VAIL Core verifies the escape sequence and gives the UI focus to VAIL Server on the Quarantined VM. Request for an escape sequence prevents the potentially compromised Vulnerable VM from spoofing a transition to capture the file's password.  
- If VAIL Core does not receive the escape sequence or receive an incorrect one from VAIL Client, then it will not give the UI focus to VAIL Server. Instead, it

will display an alert noting that the system is not in quarantined state, terminate the process, and undo the first step.

- 5) VAIL Core saves the current vulnerable state in VAIL Server.
- 6) VAIL Server requests and obtains the file decryption password from the user.  
- If the user's supplied password is incorrect, then VAIL Server will terminate the process, and VAIL Core will return UI focus to VAIL Client, and undo the first step.
- 7) VAIL Server notifies VAIL Client that the state transition is complete; it decrypts and opens the file.
- 8) VAIL Core returns UI focus to VAIL Client.

#### - Returning Back to Vulnerable State

This process overwrites the vulnerable state after it reads confidential data in plaintext. After the user closes a sensitive file, vulnerable state is rolled back to the state that was previously saved before opening the file. Since sensitive files are stored in the Quarantined VM, therefore all changes that were made to the Vulnerable VM, except those that were made to the file, will be discarded. Such security measure thwarts adversaries' attempts to leak out sensitive information locally by storing sensitive files in the Vulnerable VM after opening them in quarantined state.

When the user enters an escape sequence, VAIL begins returning from quarantined state back to vulnerable state. This time the escape sequence is used to prevent the vulnerable OS from controlling the state transition process (see Subsection B in Section V). After the user enters the escape sequence, the following steps are taken (Figure 10):

- 1) VAIL Server receives the escape sequence and passes it along to VAIL Core for verification, which in turn informs VAIL Client of the required transition.
- 2) VAIL Server waits for a static time period after the Vulnerable VM flushes writes of the disk cache to the sensitive file on the Quarantined VM (see Subsection B in Section V).
- 3) VAIL Server re-encrypts the sensitive file.
- 4) VAIL Core restores the vulnerable state that was saved before entering quarantined state. Thus, all changes that were made in the Vulnerable VM while the system was in quarantined state will be overwritten.
- 5) VAIL Core stops the DHCP client service on the Quarantined VM, and restarts the VMM's DHCP and Network Address Translation services.
- 6) VAIL Core suspends and then reboots the Vulnerable VM to reset states of its peripheral virtual devices (see Subsection B in Section V).

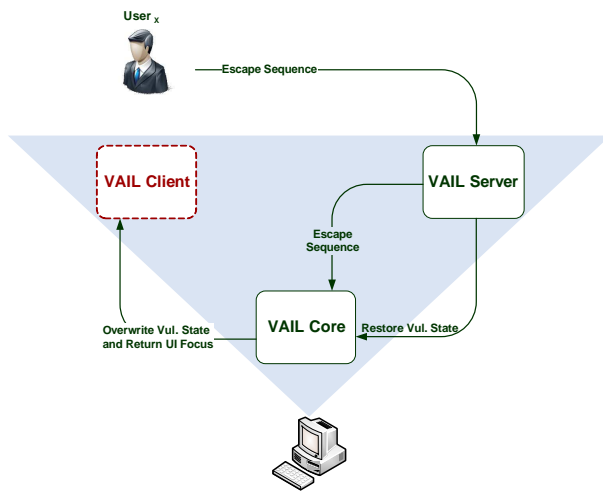


Fig. 10. VAIL process of returning back to a vulnerable state

- 7) VAIL Core resumes execution of the Vulnerable VM, re-enables its network and external storage devices drivers, and returns UI focus to VAIL Client.

## V. SECURITY EVALUATION

This section evaluates VAIL's defenses against a variety of information leakage attacks including: (1) direct information leakage attacks on sensitive files launched from: (a) the Vulnerable VM; and (b) a compromised VMM; and (2) indirect information leakage attacks exploiting covert storage and timing channels.

### A. Direct Attacks

VAIL provides its users transparent access to sensitive files that are forcibly and centrally stored encrypted in the Quarantined VM, whereas users are given the illusion of accessing them from the Vulnerable VM. However, in case adversaries knew the files locations, they may attempt to launch direct attacks against them from either the Vulnerable VM or a compromised VMM as follows.

#### - Attacks Launched from the Vulnerable VM

From the Vulnerable VM, adversaries may attempt to breach the isolation imposed on encrypted sensitive files stored in the Quarantined VM. As detailed in Subsection F in Section IV, VAIL prevents such attacks in both operational states as follows:

- While the computer system is in Vulnerable State
- VAIL Core being trusted, isolated, and highly privileged stops the DHCP client service on the Quarantined VM to prevent any communication between the Vulnerable VM and the Quarantined VM.

- VAIL prevents adversaries from capturing files decryption passwords by requesting the user to enter an escape sequence to prevent spoofing by the untrusted Vulnerable VM. In case the Vulnerable VM becomes compromised, it would ignore requesting the escape sequence and display a spoofed UI that would resemble what the user is accustomed to. Such a spoofing attack attempts to deceive the user into interacting with a malicious process instead of with VAIL Client in order to capture his/her file decryption password while the system is still in vulnerable state.

- While the computer system is in Quarantined State

- VAIL prevents adversaries from externally leaking out sensitive information from the Vulnerable VM after opening sensitive files in quarantined state. This is accomplished by VAIL Core through disabling the Vulnerable VM network and external storage devices drivers.

#### - Attacks Launched from a Compromised VMM

The VMM is an attractive target for adversaries. A compromised VMM imposes a severe threat to the Quarantined VM and its workload as adversaries acquire VMM's most privileged access level. As a result of compromising the VMM, an adversary may attempt to:

- Retrieve  $VFCK_i$  from  $CF_i$  header

VAIL confronts such an attack by two countermeasures:

-  $VFCK_i$  Encryption. By encrypting  $VFCK_i$  using a strong encryption algorithm (i.e., AES), with a long key (256-bit of  $VMCK_x$ ). Such operation strengthens data security by: (1) binding every file only with its owner; and (2) preventing illegitimate file access in case it was copied to another machine other than VAIL's Quarantined VM.

- Segregation of the Encrypted  $VFCK_i$ . VAIL splits the encrypted  $VFCK_i$  into two bitstrings. The 128 MS bits and 128 LS bits are inserted into  $FCF2\_a$  and  $FCF2\_b$  respectively. These two fields are located in the header of  $CF_i$ .

- Obtain  $VFCK_i$  before and after encryption to brute-force  $VMCK_x$ .

The objective of such an attack is to compromise all the files that were created by  $User_x$ . VAIL makes finding out  $VMCK_x$  almost impossible since the adversary would have to test  $2^{256}$  possible keys that require  $2^{256}$  attempts, which is computationally infeasible.

- Breach  $VMCK_x$  for a specific file.

Even if such an attack was successful, VAIL will prevent the adversary from compromising the user's other still encrypted sensitive files for two reasons:

- Adoption of the FLE Approach.



- VPKDV's salting and iteratively hashing the user-supplied passwords. This generates a strong unique per-file 256-bit key (i.e.,  $VMCK_x$ ). In addition, it preserves passwords' uniqueness even if a user chose identical passwords for several sensitive files, or in case multiple users chose identical passwords. Furthermore, it adds another layer of protection by hardening passwords against *dictionary-based attacks* through increasing their length. This increases the size of the search space, thereby making password-cracking computationally expensive.

### B. Indirect Attacks Exploiting Covert Channels

A covert channel is an indirect intra-machine channel that is exploited by a malicious process to transfer sensitive information with violation to the enforced security policy. Covert channels are either storage channels or timing channels [4]. They can be viewed as the worst possible indirect information leakage attack vector for their non-conventional hidden means to convey sensitive information, and the huge volume of information that could potentially be leaked, specially through covert storage channels.

A covert storage channel involves a malicious process manipulating a storage location to convey information indirectly to another storage location within a single machine. A covert timing channel allows one process to modify its usage of a shared system resource, such that, the resulting change in system response time is observed by a second process, thereby allowing it to infer extra information about sensitive data.

#### - Covert Storage Channels

Through covert storage channels adversaries may attempt to capture contents of sensitive files, and leak out devices states as follows:

- Capture contents of sensitive files

The vulnerable OS, being potentially compromised, could capture contents of sensitive files after opening them in quarantined state. It could then store them locally in the Vulnerable VM in a hidden folder in order to leak them out after returning back to vulnerable state. VAIL eliminates such covert storage channel by two countermeasures:

- Leveraging the encapsulation property of machine virtualization. Before opening a sensitive file in quarantined state, VAIL Core saves the current vulnerable state in VAIL Server. After the user closes the file, and upon his/her request, VAIL begins returning from quarantined state back to vulnerable state. VAIL Core rolls the vulnerable state back to the state that was previously captured before opening the file. Since sensitive files are stored in the Quarantined VM, therefore all changes that were made to the Vulnerable VM while the system was in quarantined state, except those that were made to the file, will be discarded (Figure 11). In addition, overwriting the entire vulnerable state extensively contributes

in confronting a wide variety of previously unknown attacks. Detailed steps were mentioned in Subsection F in Section IV.

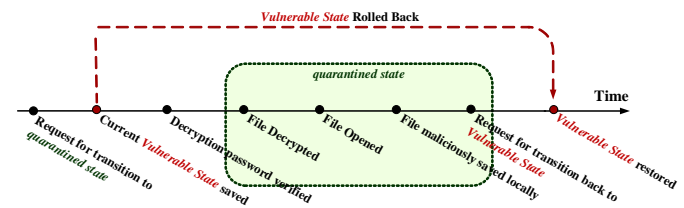


Fig. 11. VAIL's approach to eliminate covert storage channels

- Leveraging the partitioning property of machine virtualization. Since sensitive files are centrally stored encrypted in the Quarantined VM, the VMM could statically partition the host's storage space; such that, the Vulnerable VM is allocated only the minimal space required to install and run the guest OS and the required user applications.

- Leak out devices states

In a physical machine, two processes that have access to, for instance, the keyboard could leak data regarding its states (e.g., num- and caps-lock). Similarly, the Vulnerable VM could store such information while the user is entering a decryption password and/or an escape sequence to help in inferring them. For this reason, upon returning back to vulnerable state, VAIL Core suspends and then reboots the Vulnerable VM to reset states of its peripheral virtual devices (see Subsection F in Section IV).

#### - Covert Timing Channel

Through covert timing channels adversaries may attempt to:

- Control the transition timing

If the potentially compromised vulnerable OS was able to control the transition timing, then adversaries would be able to: (1) interrupt VAIL's functionality; and (2) infer information about a sensitive file after Vulnerable VM snapshot restoration. That is, after the Vulnerable VM flushes writes of the disk cache to a sensitive file on the Quarantined VM and after VAIL Core returns UI focus to the Vulnerable VM, adversaries could retrieve the system time. This would allow them to figure out the time period a user has spent editing a particular sensitive file, which would, in turn, allow them to infer its type (e.g., design document, or text document). To eliminate this covert channel, VAIL controls the state transition timing by two security measures (see Subsection F in Section IV).

- Using a predefined keyboard escape sequence. Returning back to vulnerable state is performed only at the request of the user through a predefined keyboard escape sequence.

- Adding extra delays. By adopting the approach presented in [1], VAIL imposes the timing behavior to be autonomous of

the sensitive data by adding extra delays. Such that, in each state transition, VAIL Server will wait for a static time period after flushing writes of the disk cache to a sensitive file. This would make it harder for an adversary to estimate the exact time period that was spent in viewing or editing a sensitive file.

## VI. CONCLUSION

This paper presented a novel information security approach called Virtualized Anti-Information Leakage (VAIL). Its objective was to thwart malicious software and insiders' information leakage attacks on sensitive files after decryption in potentially compromised computer systems. VAIL basically relied on leveraging machine virtualization's isolation, encapsulation, and partitioning properties to achieve its objective.

By moving VAIL's security-critical operations to an abstraction layer below that of the vulnerable operating system (OS), it isolated its security functionality from adversaries' circumvention, disabling, and subversion attacks. Through leveraging machine virtualization's encapsulation property, VAIL overcome being ad-hoc. It acquired the capability to thwart previously unknown attacks by overwriting the entire state of the untrusted potentially compromised virtual machine (VM) representing the user's personal computer without affecting the sensitive files. In addition, machine virtualization's partitioning property contributed in confining covert storage channels. VAIL benefited from advantages of file-level encryption, and overcome its drawbacks by providing its users automatic and transparent cryptographic operations.

VAIL was designed not to rely on users' commitment to security. It provided its users transparent access to sensitive files that are forcibly and centrally stored encrypted in a dedicated and isolated VM other than the VM representing the user's personal computer. In addition, a file's key is encrypted and inserted into the header of the encrypted sensitive file. Thus, a user does not need to retain any information other than his/her password. Furthermore, salting and iteratively hashing a user-supplied password preserved its uniqueness and hardened it against *dictionary-based attacks*.

At runtime, VAIL addresses *spoofing attacks* by requesting users to enter predefined escape sequences in all its critical operations. VAIL achieved the 256-bit security strength; it made *brute-forcing attacks* on a file's encryption key almost impossible since an adversary would have to test  $2^{256}$  possible keys that require  $2^{256}$  attempts, which is computationally infeasible. In addition, VAIL is directly applicable to existing commercial off-the-shelf OSs and applications, and without any special or additional required hardware.

VAIL's defenses were evaluated against a variety of information leakage attacks including: (1) *direct attacks* launched on sensitive files from an untrusted VM, and a compromised VMM; and (2) *indirect attacks* exploiting covert storage and timing channels. Based on the security evaluation, it was concluded that VAIL successfully addressed information leakage

attacks. It effectively complied with the security requirements, and met its objective. VAIL's potential users would include custodians of sensitive information in business, trade, financial, and industrial organizations.

Despite the aforementioned advantages, a limitation was identified. It arose as a result of the tireless efforts to preserve information confidentiality in a potentially compromised computer system. Opening a sensitive file terminates currently running network processes. That is, before opening a sensitive file, VAIL transits the computer system from its current *low-confidentiality state* to a *safe editing high-confidentiality state*. However, this requires disabling network device drivers of the VM that represents the user's personal computer. Consequently, network processes (e.g., downloading a file from the Internet) that may have been running before opening a sensitive file will be terminated. However, the authors believe that this limitation would not affect VAIL's usability, since, as previously mentioned, an organization's sensitive files are naturally expected to be clearly identified according to its security policy, much fewer compared to total number of business files, and with low update frequency.

## REFERENCES

- [1] J. Agat, "Transforming out Timing Leaks," in *Proc. of the 27th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL 00)*, Jan. 2000, pp. 40-53.
- [2] E. Barker, W. Barker, W. Burr, W. Polk, and M. Smid, "NIST Special Publication 800-57: Recommendation for Key Management Part 1: General (Revision 3)," National Institute of Standards and Technology (NIST), Jul. 2012.
- [3] M. Ciampa, *Security + Guide to Network Security Fundamentals*, 4th ed., Boston, Course Technology, 2012.
- [4] "Trusted Computer System Evaluation Criteria," United States Department of Defense Std., 1985.
- [5] G. Dunlap, S. King, S. Cinar, M. Basrai, and P. Chen, "ReVirt: Enabling Intrusion Analysis through Virtual-Machine Logging and Replay," in *Proc. of the 5th Symposium on Operating Systems Design and Implementation (OSDI)*, Dec. 2002. (published in a special issue of *ACM SIGOPS Operating Systems Review*), vol. 36, pp. 211-224, 2002.
- [6] G. Dunlap, D. Lucchetti, P. Chen, and M. Fetterman, "Execution Replay for Multiprocessor Virtual Machines," in *Proc. of the 4th ACM SIGPLAN/SIGOPS International Conference On Virtual Execution Environments (VEE 08)*, Mar. 2008, pp. 121-130.
- [7] Microsoft Corporation. Microsoft Support. (2012) Encrypted File System (EFS) files appear corrupted when you open them. [Online]. Available: <http://support.microsoft.com/kb/329741>
- [8] "Advanced Encryption Standard (AES)," Federal Information Processing Std. 197, Nov. 2001.
- [9] "Secure Hash Standard (SHS)," Federal Information Processing Std. 180-4, Mar. 2012.
- [10] T. Garfinkel, and M. Rosenblum, "A Virtual Machine Introspection Based Architecture for Intrusion Detection," in *Proc. of Network and Distributed Systems Security Symposium*, Feb. 2003, pp. 191-206.
- [11] W. Hagen, *Professional Xen Virtualization*, Indiana, Wiley Publishing, 2008.
- [12] J. Halderman, S. Schoen, N. Heninger, W. Clarkson, W. Paul, J. Calandrino, A. Feldman, J. Appelbaum, and E. Felten, "Lest We Remember: Cold Boot Attacks on Encryption Keys," in *Proc. of the 17th USENIX Security Symposium*, Jul. 2008, pp. 45-60.
- [13] B. Hay, and K. Nance, "Forensics Examination of Volatile System Data Using Virtual Introspection," *ACM SIGOPS Operating Systems Review*, vol. 42, no. 3, pp. 74-82, Apr. 2008.



- [14] K. Hwang, J. Dongarra, and G. Fox, "Cloud Computing: Virtualization Classes," *TechNet Magazine*, pp. 14-18, Feb. 2012.
- [15] Intel Corporation. (2013) Hardware-Assisted Virtualization Technology. [Online]. Available: <http://www.intel.com/content/www/us/en/virtualization/virtualization-technology/hardware-assist-virtualization-technology.html>
- [16] X. Jiang, X. Wang, and D. Xu, "Stealthy Malware Detection Through VMM-Based Out-of-the-Box Semantic View Reconstruction," *ACM Transactions on Information and System Security (TISSEC)*, vol. 13, no. 2, Feb. 2010.
- [17] A. Joshi, S. King, G. Dunlap, and P. Chen, "Detecting Past and Present Intrusions through Vulnerability-Specific Predicates," in *Proc. of the 20th ACM Symposium on Operating Systems Principles (SOSP 2005)*, Oct. 2005, pp. 91-104.
- [18] G. Klein, "seL4: Formal Verification of an OS Kernel," in *Proc. of the 22nd ACM Symposium on Operating Systems Principles*, Oct. 2009, pp. 207-220.
- [19] S. Kumar, U. Rawat, S. Jasra, and A. Jain, "Efficient methodology for implementation of Encrypted File System in User Space," *International Journal of Computer Science and Information Security*, vol. 3, no. 1, pp. 86-93, Jul. 2009.
- [20] J. LeVasseur, V. Uhlig, M. Chapman, P. Chubb, B. Leslie, and G. Heiser, "Pre-virtualization: soft layering for Virtual Machines," in *Proc. of the 13th IEEE Asia-Pacific Computer Systems Architecture Conference*, Aug. 2008, pp. 1-9.
- [21] L. Litty, A. Cavilla, and D. Lie, "Hypervisor Support for Identifying Covertly Executing Binaries," in *Proc. of the 17th USENIX Security Symposium*, Jul. 2008, pp. 243-258.
- [22] Microsoft Corporation. (2012) Windows Server: BitLocker Drive Encryption Overview. [Online]. Available: <http://technet.microsoft.com/en-us/library/cc732774.aspx>
- [23] A. Nguyen, N. Schear, H. Jung, A. Godiyal, S. King, and H. Nguyen, "MAVMM: Lightweight and Purpose Built VMM for Malware Analysis," in *Proc. of the 25th Annual Computer Security Applications Conference*, Dec. 2009, pp. 441-450.
- [24] R. Riley, X. Jiang, and D. Xu, "Guest-Transparent Prevention of Kernel Rootkits with VMM-based Memory Shadowing," in *Proc. of the 11th International Symposium on Recent Advances in Intrusion Detection*, Sep. 2008, pp. 1-20.
- [25] Secunia. (2013) Vulnerability Report: Xen 4.x. [Online]. Available: <http://secunia.com/advisories/product/33176/>
- [26] Secunia. (2013) Vulnerability Report: Microsoft Windows 7 [Online]. Available: <http://secunia.com/advisories/product/27467/>
- [27] A. Seshadr, M. Luk, N. Qu, and A. Perrig, "SecVisor: A Tiny Hypervisor to Provide Lifetime Kernel Code Integrity for Commodity OSes," in *Proc. of the 21st ACM SIGOPS Symposium on Operating Systems Principles*, Oct. 2007, pp. 335-350.
- [28] T. Shinagawa, H. Eiraku, K. Omote, S. Hasegawa, M. Hirano, K. Kourai, Y. Oyama, E. Kawai, K. Kono, S. Chiba, Y. Shinjo, and K. Kato, "BitVisor: A Thin Hypervisor for Enforcing I/O Device Security," in *Proc. of the ACM International Conference on Virtual Execution Environments*, Mar. 2009, pp. 121-130.
- [29] J. Smith, and R. Nair, *Virtual Machines: Versatile Platforms for Systems and Processes*, San Francisco, Morgan Kaufmann Publishers, 2005.
- [30] M. Tulloch, and Microsoft Virtualization Teams, *Understanding Microsoft Virtualization Solutions from the Desktop to the Datacenter*, 2nd ed., Washington, Microsoft Corporation, 2010.
- [31] Xen. (2006) Xen: Enterprise Grade Open Source Virtualization, A Xen White Paper V06012006 [Online]. Available: <http://www-archive.xenproject.org/files/xenWhitePaper3.2.pdf>

# Performance Analysis of Speech Quality in VoIP during Handover

M. Yousef

Electronics&Communications Dept.,  
Faculty of Eng., Zagazig Uni.,Egypt.

M. Fouad

Electronics&Communications Dept.,  
Faculty of Eng., Zagazig Uni.,Egypt.

**Abstract**— Quality of Service is a very important factor to determine the quality of a VoIP call. Different subjective and objective models exist for evaluating the speech quality in VoIP. E-model is one of the objective methods of measuring the speech quality; it considers various factors like packet loss, delay and codec impairments. The calculations of E-model are not very accurate in case of handovers – when a VoIP call moves from one wireless LAN to another. This paper conducted experimental evaluation of performance of E-model during handovers and proposes a new approach to accurately calculate the speech quality of VoIP during handovers and make MOS calculator which take the results through. A detailed description of the experimental setup and the comparison of the new approach with E-model is presented in this work.

## I. INTRODUCTION

Voice over IP services uses the traditional Internet Protocol (IP) to send the voice packets (1). It breaks the voice call into small packets that are routed over the internet. Due to the unreliable nature of the internet, these packets might get lost in the network which results in missing packets at the receiver end. As a result, the receiver would hear the speaker's sentence incomplete and may not understand it. It is very essential to monitor the quality of these voice calls to achieve user satisfaction.

To measure the speech quality various network factors like delay, packet loss, jitter etc. are considered. The measured speech quality is then mapped to a user satisfaction level. Nowadays, many people make VoIP calls when they are traveling, thus moving from one network to another. It is very important that the user experiences a good call quality when the VoIP call gets handed off from one network to another. The process of handoff consists of temporarily disconnecting from one network and then establishing a connection with the new network, this could result in dropped calls or heavy packet loss if not performed smoothly. Thus, it is very important to measure the speech quality of VoIP during handovers to achieve high user satisfaction.

## II. MEASUREMENT OF SPEECH QUALITY:

Speech quality is the measurement of user experience when a VoIP call is established (2). The measurement of speech quality is divided into two broad categories: Objective

measurement and Subjective measurement. Subjective tests are user listening tests where users are told to rate the speech quality. These tests are expensive to perform and the accuracy of speech quality rating 17 relies on the user's mood. To measure the accuracy of these subjective tests, objective methods are used. These methods are the computational methods that usually compare a good quality signal to a degraded signal (4).

### A. MEAN OPINION SCORE (MOS) – SUBJECTIVE LISTENING TEST:

Mean Opinion Score (MOS) is International Telecommunications Union Telecommunication Standardization sector (ITU-T) approved. It is a subjective listening test where the user rates the speech quality during the call.

MOS test ratings can be used to compare various codec's such as iLBC and G.711. although; MOS tests are the most reliable method of measuring the speech quality they are cumbersome to perform. They are considered as expensive tests and are quite time consuming, so it's difficult to perform them frequently.

### B. PERCEPTUAL EVALUATION OF SPEECH QUALITY (PESQ) – OBJECTIVE METHOD:

Perceptual Evaluation of Speech Quality (PESQ) is an ITU-T standard for objective measurement. It was introduced as MOS subjective tests were expensive to conduct and required a lot of time. PESQ test setup automatically maps the PESQ score to the subjective MOS score.

It takes into account two signals; one is the reference signal while the other one is the actual degraded signal. Both the signals are sent through the test that uses the PESQ algorithm and the result is a PESQ score as shown in figure 1 below.

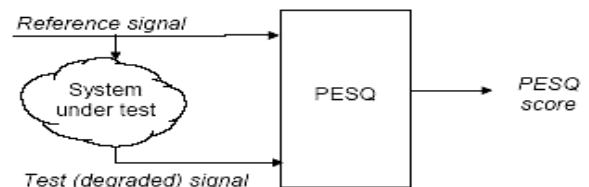


Figure 1: PESQ Testing

Major drawbacks of PESQ approach was that it did not take into account the various impairments such as acoustic echo, transmission delay etc. Also this approach cannot be used to monitor real time calls and compare codec's accurately (8).

### C. E-MODEL – OBJECTIVE APPROACHES

The E-model is transmission planning tool that provides a prediction of the expected voice quality(10), as perceived by a typical telephone user, for a complete end-to-end(i.e. mouth to ear)telephone connection under conversational conditions. The E-model takes into account a wide range of telephony-band impairments, in particular the impairment due to low-bit-rate coding devices and one-way delay, as well as the classical telephony impairments of loss, noise and echo. It is a new objective model proposed by ITU-T and it takes into account all the drawbacks of PESQ.It is a non-intrusive method of predicting the voice quality.E-model takes into account various factors that affect the speech quality and calculate a Rating factor(R-factor) that ranges between 0 – 100.the R-factor can also be converted into a MOS rating to give the MOS score

The R-factor is calculated as:

$$R_{obj} = R_0 - I_s - I_d - I_e + A \quad (1)$$

Where:

$R_0$ : Signal to Noise Ratio (S/N) at 0 dBR point

$I_s$ : Various speech impairments (e.g. Quantization noise, side tone level)

$I_d$ : Impairments that occur due to delay (e.g. absolute delay, echo)

$I_e$ : Impairments caused by the equipment (e.g. codec's, jitter, packet loss)

$A$ : Advantage factor ( $A$  is 0 for wire line and  $A$  is 5 for wireless)

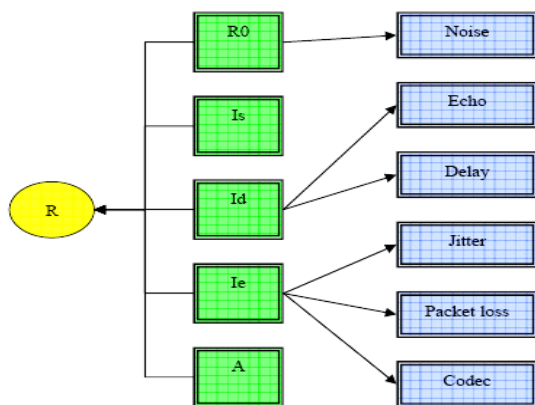


Figure 2 shows the terms of R-factor equation.

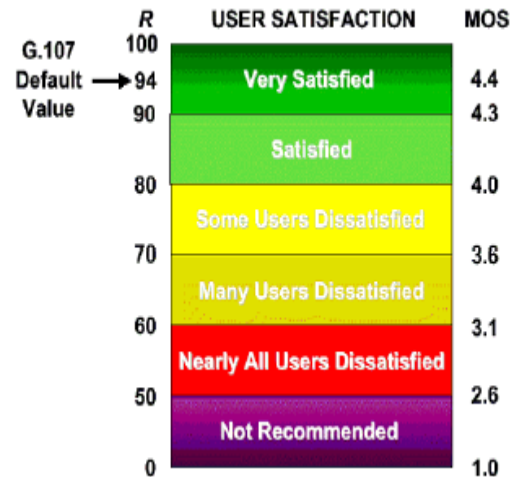


Figure 3 R-factor Rating with MOS Score Mapping and User Satisfaction Level

Based on ITU G.107 recommendation, the R- factor equation can be simplified as:

$$R\text{-factor} = 93.2 - I_d - I_e - A$$

Where:

$A$ :The Advantage factor; 0 for wire line and 5 for wireless networks.

The value of  $I_e$ , which is codec dependent impairment, is calculated as:

$$I_e = a + b \ln(1 + cP/100)$$

Where:

$P$ : The percentage packet loss and  $a$ ,  $b$  and  $c$  are codec fitting parameters.

Codec fitting parameters for iLBC (3) and G.711 (9) are summarized in table (2.1)

Table I Fitting Parameters for Codec G.711 and iLBC

Parameters	G.711	iLBC
<b>A</b>	<b>0</b>	<b>10</b>
<b>B</b>	<b>30</b>	<b>19.8</b>
<b>C</b>	<b>15</b>	<b>29.7</b>
<b>Bitrate(kb/s)/frame size(ms)</b>	<b>64/20</b>	<b>15.2/20</b>

The value of  $I_d$ , which is impairment due to delay is calculated as:

$$I_d = 0.024d + 0.11 (d - 177.3) H(d - 177.3) \quad (3.5)$$

Where:

$d$ :The total one way delay (includes serialization delay, processing delay and Propagation delay) in milliseconds.  $H(x)$  is a step function defined as:

$H(x) = 0, x < 0$  and  $H(x) = 1$  otherwise.

The drawback of E-model is that the MOS scores calculated by E-model do not correlate very well with the subjective MOS scores. Also E-model does not calculate the packet loss and delay accurately during handovers, when a VoIP call moves from one network to another.

### III. IMPACT OF HANDOFF ON VOIP

Various factors that affect the speech quality in VoIP have already been discussed, but how a speech quality of VoIP is affected by a call handover.

#### A. HANDOFF IN VOIP

Handoff is the process of transferring a call connection from one base station to another basestation in a different cell or network. Process of handoff usually takes place when a user moves around a geographical area. In a VoIP call established in a wireless network, handover happens when a user moves from one wireless network to another network.

So the handover stages can be summarized as:

Stage 1: The Mobile station (MS) communicates with the serving Base station 1.

When the MS enters the overlapping region of two networks then:

Stage 2: The MS is disconnected from the serving Base Station for a while, in this stage there is

no connection to the network.

Stage 3: A new connection is established with the target Base station 2.

#### B. QUALITY OF SERVICE DURING HANDOVER

The Quality of service of VoIP calls was subjectively calculated using the MOS tests and objectively it was calculated with E-model(7).

- 1) **SUBJECTIVE MOS TEST** : When MOS test was conducted during handover, the listeners experienced a gap (silence) for awhile (that was during the handover phase) and after handover was complete they could hear the test sentences. Some calls got dropped as the mobile user could not connect to the other wireless network. This happened due to the excessive delay during handover process; the mobile user moved out of the handover region and also did not get authenticated to the new network, leading to a dropped call.

- 2) **E-MODEL CALCULATIONS**: The MOS score was also objectively calculated using the E-model. The packets were captured using Ethereal during handover. During handover the voice call was temporarily disconnected as users did not hear anything for both G.711 and iLBC, but Ethereal showed a 0% - 0.02% packet loss for G.711 and 0.3-1.1% packet loss for iLBC codec. The drawback with E-model calculations using Wireshark tool is that it does not accurately calculate the packet loss during handover. Ethereal showed a 0% - 0.02% packet loss for G.711 and 0.3-1.1% packet loss for iLBC codec whereas the actual packet loss was much more. Thus the E-model calculations for handover scenario show a very high difference between MOS subjective and MOS objective scores.

#### C. NEW OBJECTIVE MODEL

The new objective model that I propose is based on studying the Wireshark packets during handover. The handover delay with reference to the handover stages in figure can be defined as:

The delay that occurs between the time of disconnection from BS 1 and the time of setting up the connection with the BS 2 is the handover delay.

Therefore, the handover delay can be calculated by measuring Synchronization delay, delay due to ranging information and Registration delay.

$$\text{Delay (HO)} = d(\text{sync}) + d(\text{rang}) + d(\text{reg}) \text{ -----(1)}$$

Now the packet loss during handover is also measured by calculating the packets sent during the synchronization, ranging and registration phases. The Wireshark screenshot in figure shows the packets that are being sent from 192.168.1.6 to 192.168.1.8 only but at this time user cannot listen to anything, i.e. they are the packets lost during handover.

Therefore handover packet loss will be:

$$P(h) = \text{No. of packets sent during handover phase} / \text{Total packets sent}$$

Thus from the new approach the enhanced E-model equation becomes:

$$R = 93.2 - I_{dh} - I_{eh}$$

Where:

$$I_{dh} = 0.024 * \_ + 0.11 * (\_ - 177.3) * H(\_ - 177.3)$$

Where:

$\_ = d(\text{sync}) + d(\text{rang}) + d(\text{reg}) + d(\text{net}) + \text{packetization delay} + \text{processing delay}$

Therefore:

$\_ = D(\text{ho}) + d(\text{net}) + \text{packetization delay} + \text{processing delay} \dots (2)$

From E-model  $I_e$  was as:

$I_e = a + b \ln(1 + cP/100)$

Therefore  $I_e$  for handover measurement will be:

$I_{eh} = a + b \ln(1 + c(p + P_h)/100)$

Where:

$P_h$ : The packet loss during handover

#### IV. TEST SETUP

The test setup consisted of windows machine and mobile that had a VoIP call established. The VoIP client was downloaded on both machines and was configured to use RTP ports for sending and receiving voice packets. One laptop was fixed while the user on the other laptop was mobile during handover. The MOS subjective tests were performed for VoIP call with and without handovers using both the codec's G.711 and iLBC. During each test, packets were captured using Wireshark tool

A. *VOIP CLIENT*: The VoIP client used in the project is a freely available VoIP client.

B. *MOS SCORE CALCULATOR*: The MOS score calculator that I developed is based on the new proposed E-model for handovers.

It calculates the R-factor and the corresponding MOS scores for speech samples during handovers using the new approach. The main purpose of this tool is to reduce the manual effort.

#### V. IMPLEMENTATION RESULTS

In order to calculate the speech quality of VoIP during handover, firstly MOS test was conducted for 12 participants. This subjective MOS test was performed with and without handover. Ten sample Hindi test sentences were played and participants rated each test sentence based on the quality.

##### A. SCENARIO WITHOUT HANDOVER

###### 1) MOS SCORE (OBJECTIVE) FOR G.711

E-model calculations for G.711 without handover:

Delay  $\Delta = 20\text{ms}$

Total delay  $d = \Delta + \text{packetization delay} + \text{processing delay} = 20 + 20 + 5 = 45\text{ms}$

$I_d = 0.024 * d + 0.11(d - 177.3) H(d - 177.3) = 1.08$

$I_e$  is 0 for G.711

$R = R_o - I_d - I_e = 92.12$

Therefore MOS = 4.383

###### 2) E-model calculations for iLBC without handover:

Delay  $\Delta = 30\text{ms}$

Total delay  $d = \Delta + \text{packetization delay} + \text{processing delay} = 20 + 20 + 15 = 65\text{ms}$

$I_d = 0.024 * d + 0.11(d - 177.3) H(d - 177.3) = 1.38$

$I_e$  is 10 for iLBC

$R = R_o - I_d - I_e = 81.64$

Therefore MOS = 4.084

##### B. SCENARIO WITH HANDOVER

###### 1) SUBJECTIVE MOS SCORE

When MOS test was conducted during handover, the listeners experienced a gap (a silence) for awhile (that was during the handover phase) and after handover was complete they could hear the test sentences. The MOS test was performed with 12 participants.

Some calls got dropped as the mobile user could not connect to the other wireless network. This happened due to the excessive delay during handover process; the mobile user moved out of the handover region and also did not get authenticated to the new network, leading to a dropped call.

Codec Average MOS Score

iLBC 3.143

G.711 3.311

Table Average MOS scores for G.711 and iLBC during handover

###### 2) E-MODEL CALCULATION FOR G.711

Avg one way delay  $\Delta = 20\text{ms}$

$d = 20 + 5 + 20 = 45\text{ms}$

Therefore  $I_d = 0.024 * d = 1.08$

Packet loss (P) for G.711 was 0.08%

Therefore,  $I_e = a + b \ln(1 + cP/100)$

$= 0 + 30 \ln(1 + 0.08 * 15/100) = 0.0358$

$R_{\text{factor}} = 93.2 - I_d - I_e = 91.76$

$\text{MOS (emodel)} = 1 + 0.035 * R + R(R - 60)(100 - R)^{-7} * 10^{-6}$

**MOS (E model) = 4.38**

### 3) E-MODEL CALCULATION FOR Ilbc

Packet loss = 0.12%

Delay = 30 ms

Id = 1.56

$I_e = a + b \ln(1 + C_p/100)$

$= 10 + 20 \ln(1 + 0.12 \cdot 30/100)$

$I_e = 10.707$

$R_{factor} = 93.2 - 1.56 - 10.707 = 80.93$

Therefore, **MOS (E model) = 4.059**

The results of MOS (subjective) and MOS (E-model) are summarized in figure

## VI. CALCULATIONS USING NEW APPROACH

### A. NEW E-MODEL CALCULATION FOR ILBC

Avg one way delay = 30ms

$d = 30 + 15 + 20 = 65\text{ms}$

Therefore  $I_{dh} = 0.024 \cdot d = 1.56$

Packet loss (P) for iLBC is 0.12%

Handover packet loss ( $P_h$ ) = 15.09%

Therefore,  $I_{eh} = a + b \ln(1 + c(P + P_h/100))$

$= 10 + 20 \ln(1 + 15.21 \cdot 30/100) = 44.32$

$R_{factor} = 93.2 - 44.32 = 47.32$

$MOS \text{ (New-Emodel)} = 1 + 0.035 \cdot R + R(R-60)(100-R)^7 \cdot 10^{-6}$

**MOS (New – Emodel) = 2.45**

### B. NEW E-MODEL CALCULATION FOR G.711

Avg one way delay = 20ms

$d = 20 + 5 + 20 = 45\text{ms}$

Therefore  $I_{dh} = 0.024 \cdot d = 1.08$

Packet loss (P) for G.711 was 0.08%

Handover packet loss ( $P_h$ ) = 14.75%

Therefore,  $I_{eh} = a + b \ln(1 + c(P + P_h/100))$

$= 0 + 30 \ln(1 + 14.83 \cdot 15/100) = 35.12$

$R_{factor} = 93.2 - 35.12 - 1.08 = 57$

$MOS \text{ (New-Emodel)} = 1 + 0.035 \cdot R + R(R-60)(100-R)^7 \cdot 10^{-6}$

**MOS (New – Emodel) = 2.94**

Comparison between G.711 and iLBC:

Without handover:

Delay:

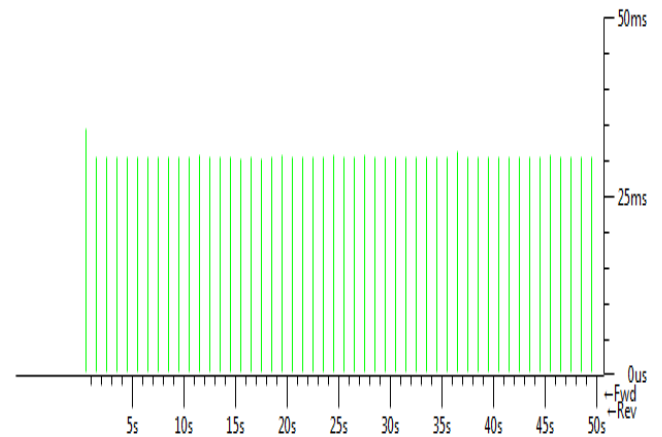
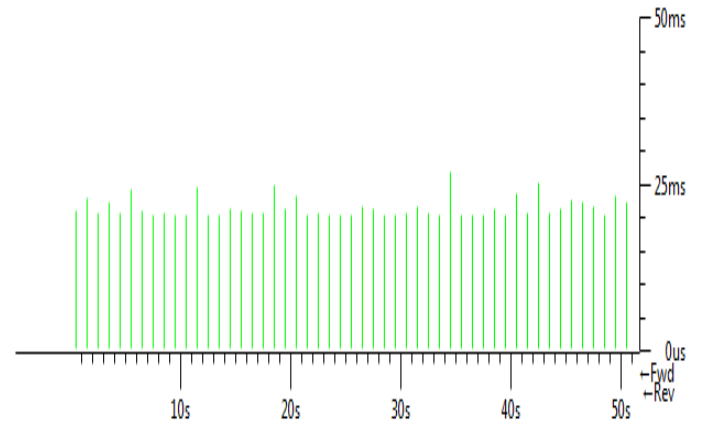


Figure 4. iLBC Delay

We note that the delay happen when use iLBC is more than what happen with G.711 as in figure(5.1) and figure(5.2).

With handover

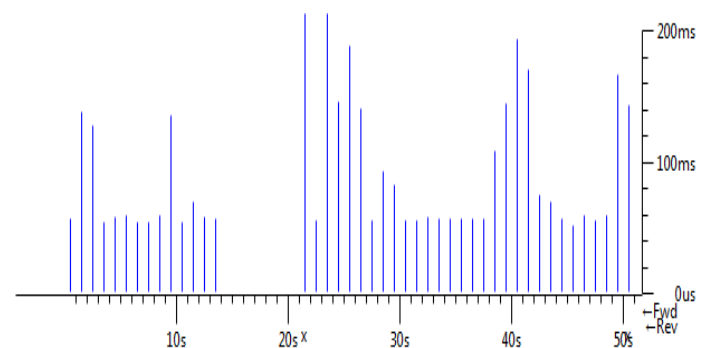


Figure5.:G.711 Delay

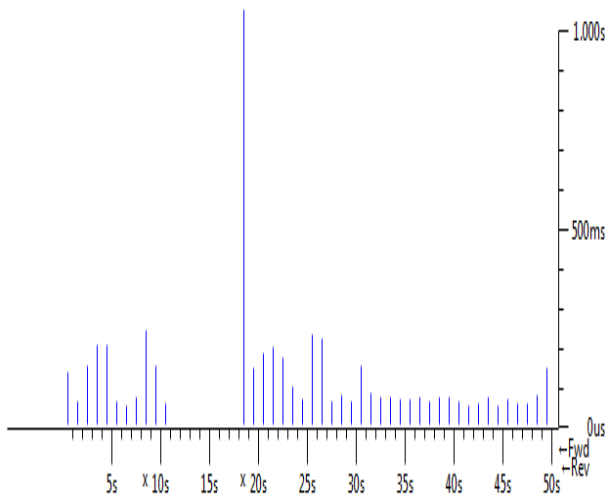


Figure 6.:iLBC Delay

Figure(5.3) and figure(5.4) show that the delay when using iLBC is much bigger than what happen with G.711 with handover.

## CONCLUSION

The results of the MOS calculator shows that the new approach maps very close to the subjective MOS scores as compared to E-model and helps to calculate the speech quality during handoff much accurately. G.711 codec is a better speech codec than iLBC codec. The speech quality for G.711 is extremely good without handoff, but during a call handoff, the speech quality does degrade but not as much as iLBC. The speech quality for iLBC is tremendously degraded during a call handover and leads to user dissatisfaction.

## REFERENCES

- [1] Jane Dudman& Gaynor Blackhouse: Voice over IP: JISC technology and Standards Watch, September 2006[2] QoE Systems. Retrieved from <http://www.qoesystems.com/audio>
- [3] S.Andersen and A.Duric, "Internet Low Bit-Rate Codec (iLBC),IETF Draft," Feb 2002.
- [4] Shengquan Wang, Zhibin Mai, Walt Magnussen, Dong Xuan and Wei Zhao:Implementation of QoS- Provisioning System for Voice over IP, In the Proceedings of theEight IEEE Real time and Embedded Technology and Applications Symposium, 2002.
- [5] Pincy C. Mehta, Sanjay Udani: Overview of Voice over IP, Technical Report MS-CIS-01-31, February 2001. Retrieved from
- [6]GIPs,iLBCRetrievedfrom[http://www.softfront.co.jp/products/lib/codecs\\_iLBC](http://www.softfront.co.jp/products/lib/codecs_iLBC)
- [7] Dong Hoi Kim and Kyungkoo Jun (October 5, 2005). The Effect of the QoS Satisfactionon the Handoff for Real-TimeTraffic in Cellular Network. In High Performance Computingand Communications (Volume 3726).
- [8] Ditech Networks whitepaper, Limitations of PESQ for Measuring Voice Quality inMobile and VoIP Networks, 2007.
- [9] Agora Labs, Speech Codecs , retrieved at <http://www.agoralabs.com/speechcodec/g711-codec.htm>
- [10] ITU-T Recommendation G.107, "The E-model, a Computational Model for use in Transmission Planning", Mar. 2005.
- [11] Design and analysis of IP multimedia subsystem/Ain Shams University 2011.



# Web Users Clustering Analysis

Hooman Rokham

Computer Engineering Department  
Shahre-e-Qods Branch, Islamic Azad University  
Shahr-e-Qods, Iran

Hale Falakshahi

Computer Engineering Department  
Science and Research Branch, Islamic Azad University  
Neyshabur, Iran

**Abstract**—As one of the most important tasks of web usage mining, web user clustering, which establishes groups of users exhibiting similar browsing patterns, provides useful knowledge to personalized web services. There are many clustering algorithm. In this paper, users' similarity is calculated then a comparative analysis of two clustering algorithms namely K-means algorithm and hierarchical algorithm is performed. Web users are clustered with these algorithms based on web user log data. Given a set of web users and their associated historical web usage data, we study their behavior characteristic and cluster them. In terms of accuracy K-means produces better results as compared to hierarchical algorithm.

**Keywords**—clustering; K-means algorithm; hierarchical algorithm

## I. INTRODUCTION

The World Wide Web has become increasingly important as a medium for commerce as well as for dissemination of information. In E-commerce, companies want to analyze the user's preferences to place advertisements, to decide their market strategy, and to provide customized guide to Web customers. In today's information based society, there is an urge for Web surfers to find the needed information from the overwhelming resources on the Internet. Web access log contains a lot of information that allows us to observe user's interest with the site. Properly exploited, this information can assist us to make improvements to the Web site, create a more effective Web site organization and to help users navigate through enormous Web documents. Therefore, data mining, which is referred to as knowledge discovery in database, has been naturally introduced to the World Wide Web. When applied to the World Wide Web, data mining is called Web mining. Web mining is categorized into three active research areas according to what part of web data is mined, of which Usage mining, also known as web-log mining, which studies user access information from logged server data in order to extract interesting usage patterns. In this context, cluster analysis can be considered as one of the most important aspects in the Web mining process for discovering meaningful groups as well as interpreting and visualizing the key behaviors exhibited by the users in each cluster. The clustering problem is about partitioning a given data set into clusters such that the data points in the same cluster are more similar to each other than points in different clusters.

In this paper, we explore the problem of user clustering and then a comparative analysis of two clustering algorithms

namely K-means algorithm and Hierarchical algorithm is performed. The performance of these clustering algorithms is compared in terms of accuracy.

The rest of the paper is organized as follows: in section 2, related work will be introduced. In section 3 method and clustering analysis will be explained. The experimental result and discuss will be introduced in section 4. Finally, section 5 concludes the paper.

## II. RELATED WORK

Several researchers have applied data mining techniques to web server logs, attempting to unlock the usage patterns of web users hidden in the log files. Data mining, which is referred to as knowledge discovery in database, has become an important research area as a consequence of the maturity of very large databases. It uses techniques from areas such as machine learning, statistics, neural networks, and genetic algorithms to extract implicit information from very large amounts of data. The goals of data mining are prediction, identification, classification, and optimization. The knowledge discovered by data mining includes association rules, sequential patterns, clusters, and classification. Garofalakis [1] gives a review of popular data mining techniques and the algorithms for discovering the Web. Reference [2] proposed a taxonomy of Web mining and identified further research issues in this field. Yu [3] examines new developments in data mining and its application to personalization in E-commerce. Reference [4] has demonstrated that web users can be clustered into meaningful groups, which help webmasters to better understand the users and therefore to provide more suitable, customized services. Mobasher, Cooley and Srivastava [5] propose a technique for capturing common user profiles based on association rule discovery and usage-based clustering. This technique directly computes overlapping clusters of URL references based on their co-occurrence patterns across user transactions.

Nasraoui and Krishnapuram [6] use unsupervised robust multi-resolution clustering techniques to discover Web user groups. Xie and Phoha [7] use belief functions to cluster Web site users. They separate users into different groups and find a common access pattern for each group of users. Xu and Liu [8] cluster web users with K-means algorithm based on web user log data; they introduced 'hits' concept, hits mean one kind of user browsing information. We can directly extract the hits of

all users who access the Web pages of a Web site during a given period of time,  $hits(i, j)$  is the count of user  $i$  accesses Web page  $j$  during a defined period of time. Count of visiting the pages is the criterion that is used for clustering. Reference [9] emphasized the need to discover similarities in users' accessing behavior with respect to the time locality of their navigational acts. In this context, they present two time-aware clustering approaches for tuning and binding the page and time visiting criteria. The two tracks of the proposed algorithms define clusters with users that show similar visiting behavior at the same time period, by varying the priority given to page or time visiting.

### III. METHOD AND CLUSTERING ANALYSIS

#### A. Calculate Users' Similarity

The method begins with preprocessing of server logs and then users' sessions are extracted. The methods of comparing similarity between users based on a criteria will be presented and user clustering will be done by two algorithms namely K-means algorithm and hierarchical algorithm. Accuracy of these two algorithm will be analyzed.

Preprocessing of web server log files is conducted to identify user sessions. Web servers often registered all the activities of users in the form of web server logs. Because of the different configurations of servers, there are several types of server logs. But normally server log files, have the same basic information, such as client IP address, time of request, requested URL, the status code of HTTP, references and more. Several preprocessing operations should perform before applying the web usage mining techniques on the web server logs. These operations in the scope of our research include data cleansing and identifying and separating users' sessions. All data in web server logs are not suitable for web usage mining. So to remove the improper data from the log file, data cleansing step is accomplished.

A user session is a set of pages seen by the user during a special visit from a website. Before applying web usage mining techniques, web server logs should be grouped into meaningful sessions. In this study pages are considered as a session that are requested in a period of time less than equal to a certain time. Two appropriate features of user which represent user's interests are 'page view frequency' and 'time of viewing the page'. After web server log's preprocessing, the amount of these features are calculated and then Cosine similarity is used to calculate the amount of similarity between each two users.

The similarity between two users can be measured by counting the number of times they access the common pages. We use the cosine similarity as the similarity measure. In this case, the measure is defined by

$$S_f(u_i, u_j) = \frac{\sum_k (\text{freq}(u_i, p_k) * \text{freq}(u_j, p_k))}{\sqrt{\sum_k \text{freq}(u_i, p_k)^2 * \sum_k \text{freq}(u_j, p_k)^2}} \quad (1)$$

Where  $S_f(u_i, u_j)$  is a similarity between user  $u_i$  and user  $u_j$  based on page view frequency.  $\text{freq}(u_i, p_j)$  shows the number of time the user  $u_i$  accessed to the page  $p_j$ .

In like manner, the similarity between two users can be measured more precisely by taking into account the actual time the users spent on viewing each web page. We use the cosine similarity as the similarity measure. In this case, the measure is defined by

$$S_t(u_i, u_j) = \frac{\sum_k (t(u_i, p_k) * t(u_j, p_k))}{\sqrt{\sum_k t(u_i, p_k)^2 * \sum_k t(u_j, p_k)^2}} \quad (2)$$

Where  $S_t(u_i, u_j)$  is a similarity between user  $u_i$  and user  $u_j$  based on time period of viewing the page.  $t(u_i, p_j)$  shows the amount of time user  $u_i$  spend on viewing the page  $p_j$ .

As a result, after using equation (1) and (2) two users- pages matrices are gained. For clustering we need to have one matrix. Ergo, equation (3) is defined by

$$w(u_i, p_j) = a \text{freq}(u_i, p_j) + b t(u_i, p_j) \quad (3)$$

Where  $w(u_i, p_j)$  is a weight given to page  $p_j$  based on user  $u_i$  features and  $a, b$  are experience values in the site. The similarity between users can also be gained based on weighting criterion by use of cosine similarity.

#### B. Clustering Algorithms

##### 1) K-Means Clustering

In this stage K-means clustering algorithm is performed. The flow of algorithm is shown as the following steps:

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

At the end of K- Means clustering stage we K clusters will be obtained that users in each cluster patterns would be most similar to each other with respect to individual preferences.

#### C. Hierarchical Clustering [10]

Hierarchical Clustering method merged or splits the similar data objects by constructing hierarchy of clusters also known as dendrogram. Hierarchical Clustering method forms clusters progressively. Hierarchical Clustering classified into two forms: Agglomerative and Divisive algorithm.

- Agglomerative hierarchical clustering is a bottom up method which starts with every single object in a

single cluster. Then, in each successive iteration, it combines the closest pair of clusters by satisfying some similarity criteria, until all of the data is in one cluster or specify by the user.

- Divisive hierarchical clustering is a top down approach. Divisive hierarchical clustering starts with one cluster that contain all data objects. Then in each successive iteration, it divide into the clusters by satisfying some similarity criteria until each data objects forms clusters its own or satisfies stopping criteria.

#### IV. EXPERIMENTAL EVALUATION

The real data set were used for this study which related to Palood dairy<sup>1</sup> products company. The result of this report is for two months (June, July 2014). The size of access log for two months is 56 MB. The information contained in the user access log of this site includes more details of users' requests so unnecessary information has been refined. The information stored for this study include: requesting IP address, date and time of the request, Parameters sent in every web address, communication method, user's browser type and operating system and page sizes. After collecting information data preprocessing was done. To implement the components of the proposed approach, the SQL Server 2008 database, Visual Studio 2010 software and RapidMiner was used. RapidMiner is a software that provides an integrated environment for machine learning, data mining, text mining, predictive analytics and business analytics.

After refining the raw data the number of distinct users for two month who visited the site was gained. The number of distinct users was 1240 and the number of valid web pages was 160.

The value of equation (3) were calculated and the similarity between users was gained based on weighting criterion by use of cosine similarity and then was given to RapidMiner software in form of Users-Pages matrix. The values of a and b in equation (3) set to 0.7 and 0.3 respectively.

Having introduced the two clustering algorithms, now turn to the discussion of these algorithms on the basis of a practical study. The experimental result of these two algorithm will be presented by using real dataset. K-Means algorithm is applied to cluster web users with different k values. The optimal precision would be gained when number of clusters (K) is set to 3, figure 1 is illustrate this result. The experimental results of both algorithm are presented in Table 1. As results show, K-Means algorithm have better accuracy in comparison with hierarchical algorithm. In like manner, K-Means algorithm take lower time.

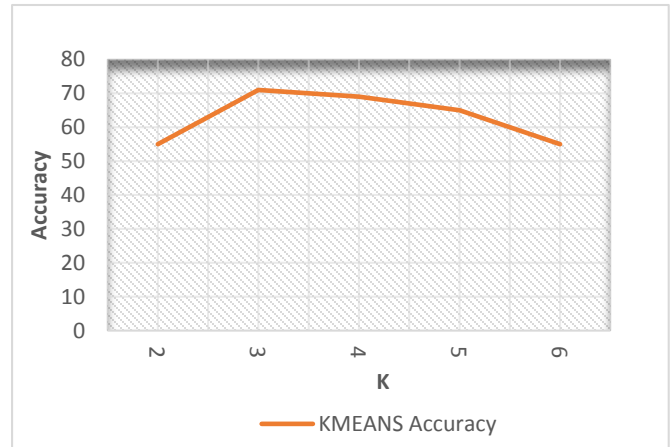


Figure 1. K-means accuracy with different k values

TABLE 1. RESULTS OF CLUSTERING

Algorithm	Number of clusters	Time	accuracy
Hierarchical	3	2.53 S	61.12%
K- Means	3	1.45 S	64.33%

#### V. CONCLUSION

In this paper, users' similarity was calculated then a comparative study has been performed on two clustering algorithms namely K-means algorithm and hierarchical clustering algorithms. Comparison was performed on real data set. Web users are clustered by these algorithms based on web user log data. K-means algorithm had a better result consider to accuracy and the time.

#### REFERENCES

- [1] M.N. Garofalakis, R. Rastogi, S. Seshadri and K. Shim, "Data mining and the Web: past, present and future," In Proceedings of the second international workshop on Web information and data management, ACM, 1999.
- [2] R. Cooley, B. Mobasher and J. Srivastava, "Web Mining: Information and Pattern Discovery on the World Wide Web," ICTAI'97, 1997.
- [3] P. Yu, "Data Mining and Personalization Technologies," Proceedings of the 6th International Conference on Database Systems for Advanced Applications, 1998.
- [4] Y. Fu, K. Sandhu and M. Shi, "Clustering of web users based on access patterns," Lecture Notes in Artificial Intelligence, vol. 1836, Springer-Verlag, Berlin, 2000, pp. 21–38.

<sup>1</sup> www.palooddairy.com

- [5] B. Mobasher, R. Cooley and J. Srivastava, "Creating Adaptive Web Sites Through Usage-based Clustering of URLs," Proceedings of the 1999 Workshop on Knowledge and Data Engineering Exchange, 1999.
- [6] O. Nasraoui and R. Krishnapuram, "An evolutionary approach to mining robust multiresolution web profiles and context sensitive URL Associations," *International Journal of Computational Intelligence and Applications*, 2(03), 339-348, 2002.
- [7] Y. Xie, V. Phoha, "Web user clustering from access log using belief function," In Proceedings of the 1st international conference on Knowledge capture, pp. 202-208ACM, 2001.
- [8] J. Xu, H. Liu, "Web user clustering analysis based on K-Means algorithm," In Information Networking and Automation (ICINA), 2010 International Conference on (Vol. 2, pp. V2-6, IEEE, 2010.
- [9] S.G. Petridou, V.A. Koutsonikola, A.I. Vakali and G.I. Papadimitriou, "Time aware web users' clustering," *Knowledge and Data Engineering, IEEE Transactions on*, 20(5), 653-667, 2008.
- [10] N. RajalingamK.Ranjini, "Hierarchical Clustering Algorithm - A Comparative Study," Volume 19– No.3, April 2011, ISSN: 0975 – 8887.

# A comparative analysis of dynamic scheduling algorithms versus the Round-Robin scheduling algorithm

Vilma Tomço

University of Tirana,

Faculty of Mathematics, Statistics and Applied Informatics

Tirana, Albania

---

Anduela Dervishi

Polytechnic University of Tirana

Faculty of Mathematical and Physical Engineering

Tirana, Albania

---

Elton Lika

Polytechnic University of Tirana

Faculty of Information Technology

Department of Informatics Engineering

“Mother Teresa” Square, Tirana, Albania

---

Igli Tafa

Polytechnic University of Tirana

Faculty of Information Technology

Department of Informatics Engineering

---

**Abstract** - Scheduling is one of the most important concepts in Operating Systems . One of the most popular algorithms is Round - Robin , which switches the processes after running the set Time Quantum (  $TQ$  ).  $TQ$  value affects the average time of Waiting and Turnaround , and the number of Context Switches (  $CS$  ). This definition can be static , which does not change , and dynamic, calculated cycle after cycle. This review builds on the study of new techniques for the determination of  $TQ$  dynamically . Initially is shown that in all cases this method is efficient and then we rank the most important techniques used . We look at how each works and the differences and their similarities . We will observe their efficiency in different parameters and the conditions in which they are effective . Finally we show that MDTQRR is most effective, minimizing the number of  $CS$  and Harm is the most effective in AVG ( Waiting and Turnaround ) Time .

**Key words** - Round-Robin , Quantum Time , Waiting Time , Turnaround Time , Context Switch, ready queue .

## I. INTRODUCTION

A process is an instance of a computer program that is executing. It includes the current values of counters, registers and variables. The difference between a program and a process

is that the program is a set of instructions while the process is an activity. Processes that are waiting to be executed by the processor are stored in a queue called the Ready Queue. The time during which the process keeps the CPU is known as Burst Time. Arrival time is the time at which the process reaches the ready queue. Waiting time is called the length of the stay of the process in the ready queue. Context switch is the number of times the CPU switches from one process to another. Turnaround time is the time from the arrival of the process to ready queue until its completion. The best algorithm would have minimum waiting, minimum turnaround time and smaller number of Context switches.

## II. SCHEDULING ALGORITHMS

### **First Come First Served (FCFS):**

In this algorithm, the first process who makes a request is selected. Although very simple, it has major shortcomings in performance compared to the other algorithms. With FCFS many short processes wait too long. If a long process takes its time, others must wait until it has finished. This effect is called the convoy.

### **Shortest Job First(SJF):**

In this algorithm short processes have priority. If two processes have the same duration, then it takes that who came first, turning into FCFS [14].

### **Shortest Remaining Time First (SRTF):**

This algorithm resembles with SJF but with some modifications. When the burst time of the processes is calculated, the remaining time to finish of the current process is taken into account.

### **Priority Scheduling Algorithm:**

It defines priorities to each of the processes and their selection is made based on the highest priority. But this method can leave a process with lower priority waiting forever. If the system is loaded it risks going to starvation. A solution could be that you increase priority of the processes that stay very long in waiting [9].

### **Round Robin Scheduling Algorithm:**

Round Robin is one of the oldest, most naive, most fair and most spread. Each of the processes has the same priority  $k$ , and is given some time, Time Quantum (TQ), and after this period has passed the process is switched. Two situations occur in this case, first, whether the given TQ is longer than its burst time the process then leaves the processor itself, secondly, if TQ is less than the burst time, then the process is switched and positioned at the end of the ready queue [13].

### **Lottery Scheduling Algorithm:**

Is an algorithm that gives each process a ticket and then randomly generates the number of the process who will take turn, and the process who has that number takes the turn.

## III. STATIC AND DYNAMIC ROUND-ROBIN

Round-Robin is an algorithm who has some advantages compared to the others. It is simple, it doesn't have interrupts and data sharing, it has no traffic and is very suited to systems who have only sequential actions.

But it's major challenge is in determining of the TQ. This is the parameter which affects in the mean Waiting and Turnaround time in the execution order [5]. One of the main subjects of study in this field is the comparison of the static time quantum, meaning that it is pre-determined and never changes in any of the cycles, and the dynamic TQ which is redefined in every cycle, renovating its inputs, which in turn are dependent on the various techniques used to determine the TQ, but mostly they have as main input the burst time and the number of processes in the ready queue. If TQ is static it causes a small number of context switches for a high TQ and a high number for context switches if TQ is small. Higher number of context switches means higher mean waiting and turnaround time and all this leads to an overhead that lowers the performance of the system [7]. So the main goal of the dynamic form is to determine the right TQ value. The different formulas are set forth below.

## IV. NEW EMBEDDED TECHNIQUES IN THE ROUND-ROBIN SCHEDULING ALGORITHM

In this field of study, the embedded techniques in the round-robin algorithm, there has have been many researches

and effort to further improve it in the last four years. Many summaries are made regarding round-robin but still no new techniques. Below will be presented the main techniques that have contributed and have the opportunity to be improved and developed in the later stages. In the way that each of the techniques will be described, there will be attached the similar techniques in defining the TQ and the publication time of this technique. It is to be emphasized that many are improvements of each other.

### *a. Shortest Remaining Burst Round-Robin(SRBRR)*

SRBRR is an improved algorithm of simple R-R giving the processor to processes with the shortest remaining burst in the form of round-robin using dynamic TQ. It performs better than RR in relation to the Waiting and Turnaround time and the number of Context Switch-s. The redefinition is done every time a new process comes in ready queue, and the Time Quantum is defined as the median of the bursts of the remaining processes in ready queue [2].

$TQ = \text{Median}(\text{remaining burst time of all processes})$

It should be emphasized that the median is calculated in an ascending sorted queue. This technique has also served as the starting point for many other researchers to develop more. And exactly at the median concept these studies were initiated.

### *b. Improved Shortest Remaining Burst Round-Robin(ISRBRR)*

A further improvement of the above technique is an algorithm named Improved-SRBRR. The difference of these two techniques is precisely the definition of TQ, this is the motive of all the techniques included in the study. The manner of operation of the algorithm is the same, when a new process arrives the variables such as burst time and the number of processes is renovated, but the difference lies in the TQ definition formula [12].

$TQ = \text{Ceil}(\sqrt{\text{median} * \text{highest burst time}})$

By looking at the formula we see that TQ is calculated from the median and the highest burst time. In this case the ceiling value of the square root is taken. Relative to the other techniques that will be studied, ISRBRR has a significant improvement over its simplest version, in all parameters such as the number of context switches, average Waiting and Turnaround time. Also this improvement is noted not only in the ascending sorted queue, but also in the descending sorted queue and the random queue. Below is its schematic:



d. Average Max Round Robin (AMRR)

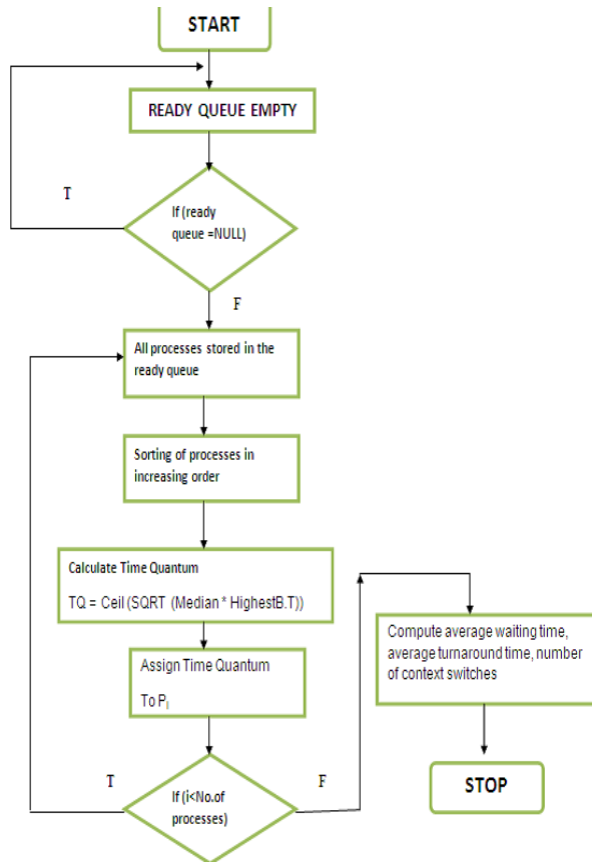
Another attempt to improve the number of Context switeches especially bringing their reduction, is also the AMRR technique. It is understood that these techniques are very similar to each other for the organization and functioning in that they all are based on quantum dynamical time. Differences are seen mainly in how TQ definition is realized and what is the object or aim of improvement, relative to the three parameters mentioned before. Another difference is also whether or not these effects are dependent on the order of the ready queue. The case in question has no impact on whether or not sorted, because of the improvement in the reduction of Context Switch-s. The formulas used in this case are two, first calculateing the average amount of time and TQ is taken as the average of this with the highest value of Burst Time [1]:

$$\text{AVG} = \text{SUM}(\text{Burst Time of all processes}) / \text{Number of processes}$$

$$\text{TQ} = (\text{AVG} + \text{MAX}(\text{BT})) / 2$$

e. Self-Adjustment Round-Robin(SARR)

A very important role in the further development of a better dynamic selection of TQ undoubtedly is played by the SARR publication. It laid the foundation of a series of new research and development that would follow. Important role and influence had the preparation in the field of mathematics of the researchers. He made two important formulas that would later be coordinated to function effectively.



Img 1. ISRBRR block diagram [12]

c. Average Mid Max Round Robin(AMMRR)

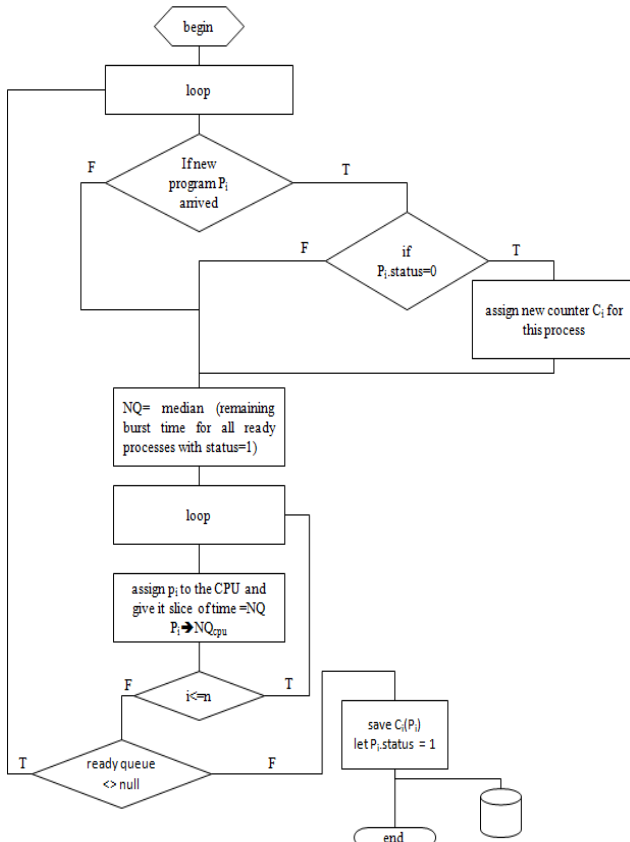
Another presented technique is AMMRR, which uses a different form of TQ calculation from the above techniques. In this case the time that is given to the process is calculated in two steps taking into account the minimum and maximum value of burst time[1].

$$\text{Mid} = (\text{Min} + \text{Max}) / 2 \quad \text{TQ} = (\text{Mid} + \text{Max}) / 2$$

So the formula has another form:

$$\text{TQ} = (\text{Min} + 3 * \text{Max}) / 4$$

In this technique too firstly ready queue is sorted and then in the form of the above techniques the TQ determination is made. But its value is determined in two steps, first calculating the average of the burst time extremes and then the average found value with the maximum of ready queue. This method has the main purpose of improving to reduce the number of switch, this time at the expense of average Waiting and Turnaround. If we take into account the last two parameters it would be more effective to use ISRBRR.



Img 2. SARR block diagram [3]

The median is again based on this concept but is a modified algorithm in the form of his choice. This is done by the odd and even numbers. The formula is as follows [3]:

$$TQ = \begin{cases} Y(N+1)/2 & \text{if } N \text{ odd} \\ \frac{1}{2} (Y(n/2)) + Y(1+N/2) & \text{if } N \text{ even} \end{cases}$$

where Y is the number of positioning among numbers listed in ascending order. But this formula can be transformed based on a simple logic that TQ can not be less than 25 being that if a process will average a TQ = 20 and to avoid situations that could lead to performance degradation we obtain a comparative TQ of 25.

$$TQ = \begin{cases} X & \text{if } X \geq 25 \\ 25 & \text{if } X < 25 \end{cases}$$

This was a method projected to maximize the CPU utilization, maximizing throughput, minimizing turnaround, waiting and response (although this is in its minimum value). Based on the simple RR that was in the time it was realized it was quite effective but with further development it was seen only as an object of study from which techniques as DQRRR, MMRR and MDTQRR were later based.

#### f. Dynamic Quantum with Readjusted Round Robin(DQRRR)

A further development of the above technique is the modified algorithm DQRRR. His largest impact was in the lowering of the context switches [4]. But again this research will be subject to further development with two other methods: MMRR and MDTQRR

$$TQ = \begin{cases} Y[(N+1)/2] & \text{if } N \text{ odd} \\ \frac{1}{2} [(Y(n/2)) + Y(1+N/2)] & \text{if } N \text{ even} \end{cases}$$

where Y is the number of positioning among numbers listed in ascending order. The above formula represents the modification made in the above-mentioned technique [3]. It has greater effectiveness if it has large number of data.

#### g. Min-Max Dispersion Measure Round Robin(MMRR)

Another variant of Round Robin is MMRR, which defines TQ time after time using minimum and maximum values of left burst time [15].

$$M = \text{MAXBT} - \text{MINBT}$$

Where : MAXBT=MAXimum Burst Time  
MINBT =MINimum Burst Time

TQ is defined which takes the newfound value from the above formula or fixed value according to the report which is given below:

$$TQ = \begin{cases} M & \text{if } M \geq 25 \\ 25 & \text{if } M < 25 \end{cases}$$

So it is an improvement of SARR. M is the space of burst times.

that other methods like SRBRR are optimized. It was built with the aim to optimize simple round robin but is worse in relation to system performance. This is based on all the determining parameters, average time of Waiting and Turnaround and number of Context Switch-s which fail to minimize and be efficient.

#### i. Multi-Dynamic Time Quantum Round Robin (MDTQRR)

This is a method which calculates the time quantum twice within a round robin cycle. This algorithm also considers the process time of arrival in the ready queue and is implemented in the algorithm. This algorithm has higher efficiency than RR and SRBRR but when is sorted in ascending order. In relation with Improved-SRBRR has very high efficiency to reduce the number of Context Switch, although an impact here has even the machine. In this algorithm are introduced two new concepts labeled MTQ (Time Quantum Media) and UTQ (Upper Quartile Time Quantum) which are calculated from the respective formulas. These two parameters are calculated once in each cycle. Formulas are given as follows:

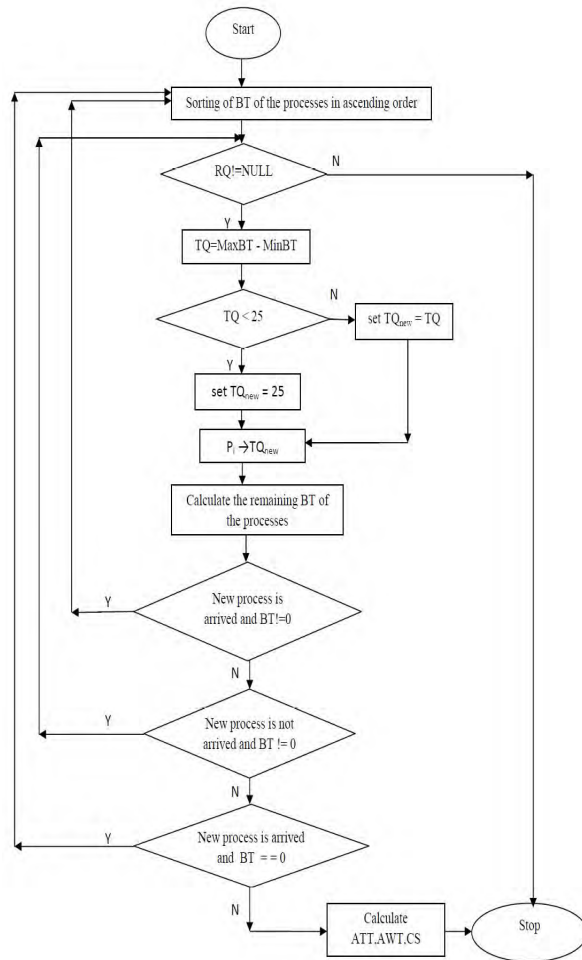
$$TQ = \begin{cases} Y(N+1)/2 & \text{if } N \text{ odd} \\ \frac{1}{2} (Y(n/2)) + Y(1+N/2) & \text{if } N \text{ even} \end{cases}$$

where Y is the number of the location in the group of numbers sorted in ascending order, and N is the number of processes. To calculate UTQ the formula is [6]:

$$UQ = \frac{3}{4}(N+1) \quad \text{where } N \text{ is the number of processes}$$

$$CRITERIA = \{ \{MTQ * m\} + \{UTQ * (N-m)\} \} / N$$

In this form the highest percentage of processes end in the startin cycles. This is seen from the chart below:



Img 3. MMRR block diagram [15]

We presented in the above block diagram how this algorithm works, but also a part of SARR given that the only added function is that of calculating the median. The benefits of this algorithm are not very high in comparison with the other algorithms under study.

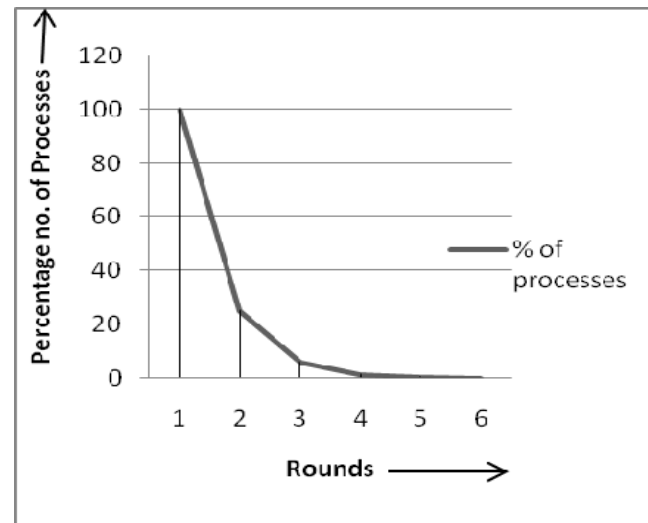
#### h. Even Odd Round Robin (EORR)

The definition of dynamic time quantum has developed another technique of determining the TQ. Two time quanta are determined on the basis of determining the position of the odd or even numbers [10]:

$$TQ1 = \text{AVG}(\text{Burst Time of Even Processes})$$

$$TQ2 = \text{AVG}(\text{Burst Time of Odd Processes})$$

Once both TQ are determined then it takes the value of one of the top two values which results lower. This form makes the burst time values close. This technique is not effective given

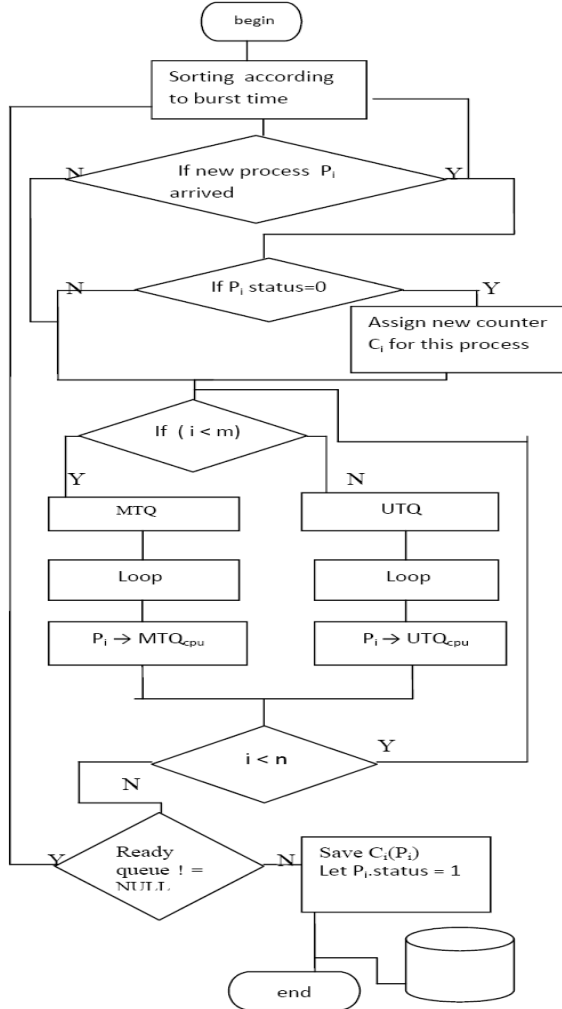


Img 4. Chart for % of processes in each round [6]

In the chart above is seen that about 75% of the processes will end in the first round and the other 25 % part does not last more than 6 rounds. This drastically decreases the context switch but also increases overhead to the operating system in many cases. Their number is given as follows:

$$Qt = [\text{sum}(Kr)] - 1$$

where Qt is the number of Context Switches, r the number of rounds and Kr the number of processes in each round. A presentation of the algorithm's block diagram is given below [11]:



Img 5. MDTQRR block diagram [11]

The complexity of this algorithm is O(n) so it is within the allowed limits.

#### j. Improved Round Robin Based on Harmonic-Arithmetic Mean (HARM)

In this algorithm is presented a new method of determining the quantum of time in each cycle based on the arithmetic average and harmonic average. In this case, judging which formula is efficient it is used one or the other. Formulas are given as follows:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

This is the formula of the arithmetic average, while the average harmonic has a slightly complex formula [15]:

$$\bar{x} = n \cdot \left( \sum_{i=1}^n \frac{1}{x_i} \right)^{-1}$$

This formula is not given often as it is difficult for many people but is quite efficient. It is a clash between three formulas of types, harmonic, arithmetic and geometric. Each is used in cases that concern us.

For a set of values in which at least two are not equal, the harmonious average is the lowest among them, the arithmetic average is the highest and geometric average is between them. So the average harmonious is affected by very low values and the opposite for the arithmetic.

So if the burst time of a new process is much smaller than the preceding process is, then its best to use the harmonic average because the new calculated average is close, it makes the average waiting time low. If a process with burst time much longer than predecessor comes, its better to use the arithmetic average instead of the harmonious.

So if burst times are heterogeneous then the harmonious average has high impact in in reducing the average waiting time (Waiting) and the average time to Turnaround.

#### V. MY APPROACH

By carefully observing and analyzing all the techniques used and their results I perceive a further development of one of the studied algorithms.

Analyzing the HARM algorithm, which has as its main advantage to minimize average waiting and turnaround time, I can say that the definition of the arithmetic mean seems wrong. If the values of burst time of the new process is much smaller then the average harmonic is used for the previously analyzed reason, but however if the value is much higher then the arithmetic average use may increase the waiting time, because the fomrula itself takes the largest value of the group. If we use another form intermediate between harmonics and arithmetic, which is the geometric average formula given below [15]:

$$\left( \prod_{i=1}^n a_i \right)^{1/n} = \sqrt[n]{a_1 a_2 \cdots a_n}$$

In one way you can also say that depends slightly on the concept of Improved-SRBRR and can lead to reducing the average waiting time. By keeping the number of context switches unchanged, since it only changes the value of TQ would be used in case of the arithmetic mean, we use the geometric, you can achieve a higher efficiency in the said parameter. But this theory must be tested to prove.

## VI. CONCLUSIONS

It should be said that many of these techniques have high sensitivity to the sorting order of burst times in the ready queue. Even many of the techniques are built on the basis of first sorting and then processing the algorithm. If a new process comes then it's sorted by the time and variables are redefined. In the three sorting forms what looks most efficient to these techniques is definitely the ascending sorted burst times list.

In all surveys made for each of the aforementioned viewed techniques the comparative etalon serving as the simplest algorithm is round-robin scheduling. And what is most important to this scheduling, it handles static forms of determining the time that is given to each of the processes and makes a comparison with the modified technique in the dynamic one. It should be noted that in all cases the dynamic technique is efficient and even in some cases it is also a drastic improvement, in the order of several times. This improved system performance is through its three main parameters, that is, the average waiting time, number of context switches and average turnaround time. So we can say that the dynamic form is very much suggested over the static one.

Surveys show that some of these techniques are designed with one primary goal, the improvement of the parameters, and specifically minimizing the number of times the processor changes the process, thus minimizing the context switches. Here we will mention which has the highest efficiency in this parameter, that is Multi-Quantum Dynamic Time Round-Robin. Although it can lead to maximum CPU use it aims to reduce the amount of CS. Unlike other techniques, here the arrival time is not zeroed but serves as input.

Two other parameters, that are average waiting time and average turnaround time, have high efficiency in two other techniques such as Improved Remaining Burst Round-Robin and HARM. Each has a very good performance in this aspect, but perhaps seeing the other parameter we could say that HARM could be slightly better compared to the first.

But in this aspect it seems there is continuous efforts for improvement and development, especially by students. Areas where there is still a problem and it is definitely pertinent to real time and embedded systems. However with these rates, given that in the last four years have been achieved so many techniques, more improvement seems to come from the work of universities.

## REFERENCES

- [1] Pallab banerjee, probal banerjee, shweta sonali dhal. Comparative Performance Analysis of Average Max Round Robin Scheduling Algorithm (AMRR) using Dynamic Time Quantum with Round Robin Scheduling Algorithm using static Time Quantum, IJITEE, ISSN: 2278-3075, Volume-1, Issue-3, August 2012.
- [2] Saroj Hiranwal. Adaptive round robin scheduling using shortest burst approach based on smart time slice, International Journal of Computer Science and Communication, vol2, No.2, July-December 011, pp.319-323.
- [3] Rami J. Matarneh. Self-Adjustment Time Quantum in Round Robin Algorithm Depending on Burst Time of Now Running Processes, American J. of Applied Sciences 6(10): 1831-1837, 2009.
- [4] H. S. Behera, Rakesh Mohanty, Debashree Nayak. A New Proposed Dynamic Quantum with Re-Adjusted Round Robin Scheduling Algorithm and Its Performance Analysis, International Journal of Computer Applications, Vol. 5, No. 5, August 2010.
- [5] Abbas Noon Ali Kalakech Seifedine Kadry. A New Round Robin Based Scheduling Algorithm for Operating Systems-Dynamic Quantum Using the Mean Average, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3, No. 1, May 2011.
- [6] S. M. Mostafa, S. Z. Rida and S. H. Hamad, "Finding Time Quantum of Round Robin CPU Scheduling Algorithm in General Computing Systems using Integer Programming", IJRRAS 5 (1), pp.64-71, October 2010.
- [7] C. Yaashuwanth, Dr. R. Ramesh. "A New Scheduling Algorithms for Real Time Tasks", (IJCSIS) International Journal of Computer Science and Information Security, Vol.6, No.2, 2009.
- [8] Tanebaun, A.S. 2008. Modern Operating Systems. 3rd Edn., Prentice Hall, ISBN: 13:9780136006633, pp: 1104.
- [9] Biju K Raveendran, Sundar Bala Subramaniam, S.Gurunarayanan, "Evaluation of Priority Based Realtime Scheduling Algorithms: Choices and Tradeoffs", SAC'08, March 16- 20, 2008.
- [10] Back, D.S., K. Pyun, S.M. Lee, J. Cho and N. Kim, 2007. A hierarchical deficit round-robin scheduling algorithm for a high level of fair service. Proceedings of the International Symposium on Information Technology Convergence, Nov. 23-24, IEEE Computer Society, Washington DC., USA.
- [11] Finley, D., J.R. Ramos, V. Rego and J. Sang, 2009. A fast computational algorithm for simulating round-robin service. J. Simulat., 3: 29-39. DOI: 10.1057/jos.2008.10
- [12] Prof. Rakesh Mohanty, Manas Ranjan Das, M.Lakshimi prasanna, Sudhashree Students. Design and Performance Evaluation of a New Proposed Fittest Job First Dynamic Round Robin (FJFDRR) Scheduling Algorithm, International Journal of Computer Information Systems, vol. 2, No. 2, 2011.
- [13] Rasmus V. Rasmussen and Michael A. Trick. Round robin scheduling a survey. European Journal of Operational Research, 188(3):617-636, August 2008.
- [14] E.O. Oyetunji, A. E. Oluleye, "Performance Assessment of Some CPU Scheduling Algorithms", Research Journal of Information Technology, 1(1): pp 22-26, 2009.
- [15] Burst Round Robin as a Proportional-Share Scheduling Algorithm Tarek Helmy, Abdelkader Dekdouk, College of





## IJCSIS AUTHORS' & REVIEWERS' LIST

Assist Prof (Dr.) M. Emre Celebi, Louisiana State University in Shreveport, USA  
Dr. Lam Hong Lee, Universiti Tunku Abdul Rahman, Malaysia  
Dr. Shimon K. Modi, Director of Research BSPA Labs, Purdue University, USA  
Dr. Jianguo Ding, Norwegian University of Science and Technology (NTNU), Norway  
Assoc. Prof. N. Jaisankar, VIT University, Vellore, Tamilnadu, India  
Dr. Amogh Kavimandan, The Mathworks Inc., USA  
Dr. Ramasamy Mariappan, Vinayaka Missions University, India  
Dr. Yong Li, School of Electronic and Information Engineering, Beijing Jiaotong University, P.R. China  
Assist. Prof. Sugam Sharma, NIET, India / Iowa State University, USA  
Dr. Jorge A. Ruiz-Vanoye, Universidad Autónoma del Estado de Morelos, Mexico  
Dr. Neeraj Kumar, SMVD University, Katra (J&K), India  
Dr Genge Bela, "Petru Maior" University of Targu Mures, Romania  
Dr. Junjie Peng, Shanghai University, P. R. China  
Dr. Ilhem LENGILIZ, HANA Group - CRISTAL Laboratory, Tunisia  
Prof. Dr. Durgesh Kumar Mishra, Acropolis Institute of Technology and Research, Indore, MP, India  
Dr. Jorge L. Hernández-Ardieta, University Carlos III of Madrid, Spain  
Prof. Dr.C.Suresh Gnana Dhas, Anna University, India  
Dr Li Fang, Nanyang Technological University, Singapore  
Prof. Pijush Biswas, RCC Institute of Information Technology, India  
Dr. Siddhivinayak Kulkarni, University of Ballarat, Ballarat, Victoria, Australia  
Dr. A. Arul Lawrence, Royal College of Engineering & Technology, India  
Dr. Wongyos Keardsri, Chulalongkorn University, Bangkok, Thailand  
Dr. Somesh Kumar Dewangan, CSVTU Bhilai (C.G.)/ Dimat Raipur, India  
Dr. Hayder N. Jasem, University Putra Malaysia, Malaysia  
Dr. A.V.Senthil Kumar, C. M. S. College of Science and Commerce, India  
Dr. R. S. Karthik, C. M. S. College of Science and Commerce, India  
Dr. P. Vasant, University Technology Petronas, Malaysia  
Dr. Wong Kok Seng, Soongsil University, Seoul, South Korea  
Dr. Praveen Ranjan Srivastava, BITS PILANI, India  
Dr. Kong Sang Kelvin, Leong, The Hong Kong Polytechnic University, Hong Kong  
Dr. Mohd Nazri Ismail, Universiti Kuala Lumpur, Malaysia  
Dr. Rami J. Matarneh, Al-isra Private University, Amman, Jordan  
Dr Ojesanmi Olusegun Ayodeji, Ajayi Crowther University, Oyo, Nigeria  
Dr. Riktesh Srivastava, Skyline University, UAE  
Dr. Oras F. Baker, UCSI University - Kuala Lumpur, Malaysia  
Dr. Ahmed S. Ghiduk, Faculty of Science, Beni-Suef University, Egypt  
and Department of Computer science, Taif University, Saudi Arabia  
Dr. Tirthankar Gayen, IIT Kharagpur, India  
Dr. Huei-Ru Tseng, National Chiao Tung University, Taiwan

Prof. Ning Xu, Wuhan University of Technology, China  
Dr Mohammed Salem Binwahlan, Hadhramout University of Science and Technology, Yemen  
& Universiti Teknologi Malaysia, Malaysia.  
Dr. Aruna Ranganath, Bhoj Reddy Engineering College for Women, India  
Dr. Hafeezullah Amin, Institute of Information Technology, KUST, Kohat, Pakistan  
Prof. Syed S. Rizvi, University of Bridgeport, USA  
Dr. Shahbaz Pervez Chattha, University of Engineering and Technology Taxila, Pakistan  
Dr. Shishir Kumar, Jaypee University of Information Technology, Wakanaghat (HP), India  
Dr. Shahid Mumtaz, Portugal Telecommunication, Instituto de Telecomunicações (IT) , Aveiro, Portugal  
Dr. Rajesh K Shukla, Corporate Institute of Science & Technology Bhopal M P  
Dr. Poonam Garg, Institute of Management Technology, India  
Dr. S. Mehta, Inha University, Korea  
Dr. Dilip Kumar S.M, University Visvesvaraya College of Engineering (UVCE), Bangalore University, Bangalore  
Prof. Malik Sikander Hayat Khiyal, Fatima Jinnah Women University, Rawalpindi, Pakistan  
Dr. Virendra Gomase , Department of Bioinformatics, Padmashree Dr. D.Y. Patil University  
Dr. Irraivan Elamvazuthi, University Technology PETRONAS, Malaysia  
Dr. Saqib Saeed, University of Siegen, Germany  
Dr. Pavan Kumar Gorakavi, IPMA-USA [YC]  
Dr. Ahmed Nabih Zaki Rashed, Menoufia University, Egypt  
Prof. Shishir K. Shandilya, Rukmani Devi Institute of Science & Technology, India  
Dr. J. Komala Lakshmi, SNR Sons College, Computer Science, India  
Dr. Muhammad Sohail, KUST, Pakistan  
Dr. Manjaiah D.H, Mangalore University, India  
Dr. S Santhosh Baboo, D.G.Vaishnav College, Chennai, India  
Prof. Dr. Mokhtar Beldjehem, Sainte-Anne University, Halifax, NS, Canada  
Dr. Deepak Laxmi Narasimha, Faculty of Computer Science and Information Technology, University of Malaya, Malaysia  
Prof. Dr. Arunkumar Thangavelu, Vellore Institute Of Technology, India  
Dr. M. Azath, Anna University, India  
Dr. Md. Rabiul Islam, Rajshahi University of Engineering & Technology (RUET), Bangladesh  
Dr. Aos Alaa Zaidan Ansaef, Multimedia University, Malaysia  
Dr Suresh Jain, Professor (on leave), Institute of Engineering & Technology, Devi Ahilya University, Indore (MP) India,  
Dr. Mohammed M. Kadhum, Universiti Utara Malaysia  
Dr. Hanumanthappa. J. University of Mysore, India  
Dr. Syed Ishtiaque Ahmed, Bangladesh University of Engineering and Technology (BUET)  
Dr Akinola Solomon Olalekan, University of Ibadan, Ibadan, Nigeria  
Dr. Santosh K. Pandey, Department of Information Technology, The Institute of Chartered Accountants of India  
Dr. P. Vasant, Power Control Optimization, Malaysia  
Dr. Petr Ivankov, Automatika - S, Russian Federation

Dr. Utkarsh Seetha, Data Infosys Limited, India  
Mrs. Priti Maheshwary, Maulana Azad National Institute of Technology, Bhopal  
Dr. (Mrs) Padmavathi Ganapathi, Avinashilingam University for Women, Coimbatore  
Assist. Prof. A. Neela madheswari, Anna university, India  
Prof. Ganesan Ramachandra Rao, PSG College of Arts and Science, India  
Mr. Kamanashis Biswas, Daffodil International University, Bangladesh  
Dr. Atul Gonsai, Saurashtra University, Gujarat, India  
Mr. Angkoon Phinyomark, Prince of Songkla University, Thailand  
Mrs. G. Nalini Priya, Anna University, Chennai  
Dr. P. Subashini, Avinashilingam University for Women, India  
Assoc. Prof. Vijay Kumar Chakka, Dhirubhai Ambani IICT, Gandhinagar ,Gujarat  
Mr Jitendra Agrawal, : Rajiv Gandhi Proudhyogiki Vishwavidyalaya, Bhopal  
Mr. Vishal Goyal, Department of Computer Science, Punjabi University, India  
Dr. R. Baskaran, Department of Computer Science and Engineering, Anna University, Chennai  
Assist. Prof, Kanwalvir Singh Dhindsa, B.B.S.B.Engg.College, Fatehgarh Sahib (Punjab), India  
Dr. Jamal Ahmad Dargham, School of Engineering and Information Technology, Universiti Malaysia Sabah  
Mr. Nitin Bhatia, DAV College, India  
Dr. Dhavachelvan Ponnurangam, Pondicherry Central University, India  
Dr. Mohd Faizal Abdollah, University of Technical Malaysia, Malaysia  
Assist. Prof. Sonal Chawla, Panjab University, India  
Dr. Abdul Wahid, AKG Engg. College, Ghaziabad, India  
Mr. Arash Habibi Lashkari, University of Malaya (UM), Malaysia  
Mr. Md. Rajibul Islam, Ibnu Sina Institute, University Technology Malaysia  
Professor Dr. Sabu M. Thampi, .B.S Institute of Technology for Women, Kerala University, India  
Mr. Noor Muhammed Nayeem, Université Lumière Lyon 2, 69007 Lyon, France  
Dr. Himanshu Aggarwal, Department of Computer Engineering, Punjabi University, India  
Prof R. Naidoo, Dept of Mathematics/Center for Advanced Computer Modelling, Durban University of Technology, Durban,South Africa  
Prof. Mydhili K Nair, M S Ramaiah Institute of Technology(M.S.R.I.T), Affiliated to Visweswaraiah Technological University, Bangalore, India  
M. Prabu, Adhiyamaan College of Engineering/Anna University, India  
Mr. Swakkhar Shatabda, Department of Computer Science and Engineering, United International University, Bangladesh  
Dr. Abdur Rashid Khan, ICIT, Gomal University, Dera Ismail Khan, Pakistan  
Mr. H. Abdul Shabeer, I-Nautix Technologies,Chennai, India  
Dr. M. Aramudhan, Perunthalaivar Kamarajar Institute of Engineering and Technology, India  
Dr. M. P. Thapliyal, Department of Computer Science, HNB Garhwal University (Central University), India  
Dr. Shahaboddin Shamshirband, Islamic Azad University, Iran  
Mr. Zeashan Hameed Khan, : Université de Grenoble, France  
Prof. Anil K Ahlawat, Ajay Kumar Garg Engineering College, Ghaziabad, UP Technical University, Lucknow  
Mr. Longe Olumide Babatope, University Of Ibadan, Nigeria  
Associate Prof. Raman Maini, University College of Engineering, Punjabi University, India

Dr. Maslin Masrom, University Technology Malaysia, Malaysia  
Sudipta Chattopadhyay, Jadavpur University, Kolkata, India  
Dr. Dang Tuan NGUYEN, University of Information Technology, Vietnam National University - Ho Chi Minh City  
Dr. Mary Lourde R., BITS-PILANI Dubai , UAE  
Dr. Abdul Aziz, University of Central Punjab, Pakistan  
Mr. Karan Singh, Gautam Budtha University, India  
Mr. Avinash Pokhriyal, Uttar Pradesh Technical University, Lucknow, India  
Associate Prof Dr Zuraini Ismail, University Technology Malaysia, Malaysia  
Assistant Prof. Yasser M. Alginahi, College of Computer Science and Engineering, Taibah University, Madinah Munawwarah, KSA  
Mr. Dakshina Ranjan Kisku, West Bengal University of Technology, India  
Mr. Raman Kumar, Dr B R Ambedkar National Institute of Technology, Jalandhar, Punjab, India  
Associate Prof. Samir B. Patel, Institute of Technology, Nirma University, India  
Dr. M.Munir Ahamed Rabbani, B. S. Abdur Rahman University, India  
Asst. Prof. Koushik Majumder, West Bengal University of Technology, India  
Dr. Alex Pappachen James, Queensland Micro-nanotechnology center, Griffith University, Australia  
Assistant Prof. S. Hariharan, B.S. Abdur Rahman University, India  
Asst Prof. Jasmine. K. S, R.V.College of Engineering, India  
Mr Naushad Ali Mamode Khan, Ministry of Education and Human Resources, Mauritius  
Prof. Mahesh Goyani, G H Patel Collge of Engg. & Tech, V.V.N, Anand, Gujarat, India  
Dr. Mana Mohammed, University of Tlemcen, Algeria  
Prof. Jatinder Singh, Universal Institutiion of Engg. & Tech. CHD, India  
Mrs. M. Anandhavalli Gauthaman, Sikkim Manipal Institute of Technology, Majitar, East Sikkim  
Dr. Bin Guo, Institute Telecom SudParis, France  
Mrs. Maleika Mehr Nigar Mohamed Heenaye-Mamode Khan, University of Mauritius  
Prof. Pijush Biswas, RCC Institute of Information Technology, India  
Mr. V. Bala Dhandayuthapani, Mekelle University, Ethiopia  
Dr. Irfan Syamsuddin, State Polytechnic of Ujung Pandang, Indonesia  
Mr. Kavi Kumar Khedo, University of Mauritius, Mauritius  
Mr. Ravi Chandiran, Zagro Singapore Pte Ltd. Singapore  
Mr. Milindkumar V. Sarode, Jawaharlal Darda Institute of Engineering and Technology, India  
Dr. Shamimul Qamar, KSJ Institute of Engineering & Technology, India  
Dr. C. Arun, Anna University, India  
Assist. Prof. M.N.Birje, Basaveshwar Engineering College, India  
Prof. Hamid Reza Naji, Department of Computer Enigneering, Shahid Beheshti University, Tehran, Iran  
Assist. Prof. Debasis Giri, Department of Computer Science and Engineering, Haldia Institute of Technology  
Subhabrata Barman, Haldia Institute of Technology, West Bengal  
Mr. M. I. Lali, COMSATS Institute of Information Technology, Islamabad, Pakistan  
Dr. Feroz Khan, Central Institute of Medicinal and Aromatic Plants, Lucknow, India  
Mr. R. Nagendran, Institute of Technology, Coimbatore, Tamilnadu, India  
Mr. Amnach Khawne, King Mongkut's Institute of Technology Ladkrabang, Ladkrabang, Bangkok, Thailand

Dr. P. Chakrabarti, Sir Padampat Singhanian University, Udaipur, India  
Mr. Nafiz Imtiaz Bin Hamid, Islamic University of Technology (IUT), Bangladesh.  
Shahab-A. Shamshirband, Islamic Azad University, Chalous, Iran  
Prof. B. Priestly Shan, Anna Univeristy, Tamilnadu, India  
Venkatramreddy Velma, Dept. of Bioinformatics, University of Mississippi Medical Center, Jackson MS USA  
Akshi Kumar, Dept. of Computer Engineering, Delhi Technological University, India  
Dr. Umesh Kumar Singh, Vikram University, Ujjain, India  
Mr. Serguei A. Mokhov, Concordia University, Canada  
Mr. Lai Khin Wee, Universiti Teknologi Malaysia, Malaysia  
Dr. Awadhesh Kumar Sharma, Madan Mohan Malviya Engineering College, India  
Mr. Syed R. Rizvi, Analytical Services & Materials, Inc., USA  
Dr. S. Karthik, SNS College of Technology, India  
Mr. Syed Qasim Bukhari, CIMET (Universidad de Granada), Spain  
Mr. A.D.Potgantwar, Pune University, India  
Dr. Himanshu Aggarwal, Punjabi University, India  
Mr. Rajesh Ramachandran, Naipunya Institute of Management and Information Technology, India  
Dr. K.L. Shunmuganathan, R.M.K Engg College, Kavaraipettai, Chennai  
Dr. Prasant Kumar Pattnaik, KIST, India.  
Dr. Ch. Aswani Kumar, VIT University, India  
Mr. Ijaz Ali Shoukat, King Saud University, Riyadh KSA  
Mr. Arun Kumar, Sir Padam Pat Singhanian University, Udaipur, Rajasthan  
Mr. Muhammad Imran Khan, Universiti Teknologi PETRONAS, Malaysia  
Dr. Natarajan Meghanathan, Jackson State University, Jackson, MS, USA  
Mr. Mohd Zaki Bin Mas'ud, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia  
Prof. Dr. R. Geetharamani, Dept. of Computer Science and Eng., Rajalakshmi Engineering College, India  
Dr. Smita Rajpal, Institute of Technology and Management, Gurgaon, India  
Dr. S. Abdul Khader Jilani, University of Tabuk, Tabuk, Saudi Arabia  
Mr. Syed Jamal Haider Zaidi, Bahria University, Pakistan  
Dr. N. Devarajan, Government College of Technology, Coimbatore, Tamilnadu, INDIA  
Mr. R. Jagadeesh Kannan, RMK Engineering College, India  
Mr. Deo Prakash, Shri Mata Vaishno Devi University, India  
Mr. Mohammad Abu Naser, Dept. of EEE, IUT, Gazipur, Bangladesh  
Assist. Prof. Prasun Ghosal, Bengal Engineering and Science University, India  
Mr. Md. Golam Kaosar, School of Engineering and Science, Victoria University, Melbourne City, Australia  
Mr. R. Mahammad Shafi, Madanapalle Institute of Technology & Science, India  
Dr. F.Sagayaraj Francis, Pondicherry Engineering College, India  
Dr. Ajay Goel, HIET, Kaithal, India  
Mr. Nayak Sunil Kashibarao, Bahirji Smarak Mahavidyalaya, India  
Mr. Suhas J Manangi, Microsoft India  
Dr. Kalyankar N. V., Yeshwant Mahavidyalaya, Nanded, India  
Dr. K.D. Verma, S.V. College of Post graduate studies & Research, India  
Dr. Amjad Rehman, University Technology Malaysia, Malaysia

Mr. Rachit Garg, L K College, Jalandhar, Punjab

Mr. J. William, M.A.M college of Engineering, Trichy, Tamilnadu, India

Prof. Jue-Sam Chou, Nanhua University, College of Science and Technology, Taiwan

Dr. Thorat S.B., Institute of Technology and Management, India

Mr. Ajay Prasad, Sir Padampat Singhania University, Udaipur, India

Dr. Kamaljit I. Lakhtaria, Atmiya Institute of Technology & Science, India

Mr. Syed Rafiul Hussain, Ahsanullah University of Science and Technology, Bangladesh

Mrs Fazeela Tunnisa, Najran University, Kingdom of Saudi Arabia

Mrs Kavita Taneja, Maharishi Markandeshwar University, Haryana, India

Mr. Maniyar Shiraz Ahmed, Najran University, Najran, KSA

Mr. Anand Kumar, AMC Engineering College, Bangalore

Dr. Rakesh Chandra Gangwar, Beant College of Engg. & Tech., Gurdaspur (Punjab) India

Dr. V V Rama Prasad, Sree Vidyanikethan Engineering College, India

Assist. Prof. Neetesh Kumar Gupta, Technocrats Institute of Technology, Bhopal (M.P.), India

Mr. Ashish Seth, Uttar Pradesh Technical University, Lucknow, UP India

Dr. V V S S S Balaram, Sreenidhi Institute of Science and Technology, India

Mr Rahul Bhatia, Lingaya's Institute of Management and Technology, India

Prof. Niranjana Reddy, P, KITS, Warangal, India

Prof. Rakesh. Lingappa, Vijetha Institute of Technology, Bangalore, India

Dr. Mohammed Ali Hussain, Nimra College of Engineering & Technology, Vijayawada, A.P., India

Dr. A.Srinivasan, MNM Jain Engineering College, Rajiv Gandhi Salai, Thorapakkam, Chennai

Mr. Rakesh Kumar, M.M. University, Mullana, Ambala, India

Dr. Lena Khaled, Zarqa Private University, Aman, Jordan

Ms. Supriya Kapoor, Patni/Lingaya's Institute of Management and Tech., India

Dr. Tossapon Boongoen, Aberystwyth University, UK

Dr. Bilal Alatas, Firat University, Turkey

Assist. Prof. Jyoti Praaksh Singh, Academy of Technology, India

Dr. Ritu Soni, GNG College, India

Dr. Mahendra Kumar, Sagar Institute of Research & Technology, Bhopal, India.

Dr. Binod Kumar, Lakshmi Narayan College of Tech.(LNCT) Bhopal India

Dr. Muzhir Shaban Al-Ani, Amman Arab University Amman – Jordan

Dr. T.C. Manjunath, ATRIA Institute of Tech, India

Mr. Muhammad Zakarya, COMSATS Institute of Information Technology (CIIT), Pakistan

Assist. Prof. Harmunish Taneja, M. M. University, India

Dr. Chitra Dhawale, SICSR, Model Colony, Pune, India

Mrs Sankari Muthukaruppan, Nehru Institute of Engineering and Technology, Anna University, India

Mr. Aaqif Afzaal Abbasi, National University Of Sciences And Technology, Islamabad

Prof. Ashutosh Kumar Dubey, Trinity Institute of Technology and Research Bhopal, India

Mr. G. Appasami, Dr. Pauls Engineering College, India

Mr. M Yasin, National University of Science and Tech, Karachi (NUST), Pakistan

Mr. Yaser Miaji, University Utara Malaysia, Malaysia

Mr. Shah Ahsanul Haque, International Islamic University Chittagong (IIUC), Bangladesh



Prof. (Dr) Syed Abdul Sattar, Royal Institute of Technology & Science, India  
Dr. S. Sasikumar, Roever Engineering College  
Assist. Prof. Monit Kapoor, Maharishi Markandeshwar University, India  
Mr. Nwaocha Vivian O, National Open University of Nigeria  
Dr. M. S. Vijaya, GR Govindarajulu School of Applied Computer Technology, India  
Assist. Prof. Chakresh Kumar, Manav Rachna International University, India  
Mr. Kunal Chadha , R&D Software Engineer, Gemalto, Singapore  
Mr. Mueen Uddin, Universiti Teknologi Malaysia, UTM , Malaysia  
Dr. Dhuha Basheer abdullah, Mosul university, Iraq  
Mr. S. Audithan, Annamalai University, India  
Prof. Vijay K Chaudhari, Technocrats Institute of Technology , India  
Associate Prof. Mohd Ilyas Khan, Technocrats Institute of Technology , India  
Dr. Vu Thanh Nguyen, University of Information Technology, HoChiMinh City, VietNam  
Assist. Prof. Anand Sharma, MITS, Lakshmangarh, Sikar, Rajasthan, India  
Prof. T V Narayana Rao, HITAM Engineering college, Hyderabad  
Mr. Deepak Gour, Sir Padampat Singhanian University, India  
Assist. Prof. Amutharaj Joyson, Kalasalingam University, India  
Mr. Ali Balador, Islamic Azad University, Iran  
Mr. Mohit Jain, Maharaja Surajmal Institute of Technology, India  
Mr. Dilip Kumar Sharma, GLA Institute of Technology & Management, India  
Dr. Debojyoti Mitra, Sir padampat Singhanian University, India  
Dr. Ali Dehghantanha, Asia-Pacific University College of Technology and Innovation, Malaysia  
Mr. Zhao Zhang, City University of Hong Kong, China  
Prof. S.P. Setty, A.U. College of Engineering, India  
Prof. Patel Rakeshkumar Kantilal, Sankalchand Patel College of Engineering, India  
Mr. Biswajit Bhowmik, Bengal College of Engineering & Technology, India  
Mr. Manoj Gupta, Apex Institute of Engineering & Technology, India  
Assist. Prof. Ajay Sharma, Raj Kumar Goel Institute Of Technology, India  
Assist. Prof. Ramveer Singh, Raj Kumar Goel Institute of Technology, India  
Dr. Hanan Elazhary, Electronics Research Institute, Egypt  
Dr. Hosam I. Faiq, USM, Malaysia  
Prof. Dipti D. Patil, MAEER's MIT College of Engg. & Tech, Pune, India  
Assist. Prof. Devendra Chack, BCT Kumaon engineering College Dwarahat Almora, India  
Prof. Manpreet Singh, M. M. Engg. College, M. M. University, India  
Assist. Prof. M. Sadiq ali Khan, University of Karachi, Pakistan  
Mr. Prasad S. Halgaonkar, MIT - College of Engineering, Pune, India  
Dr. Imran Ghani, Universiti Teknologi Malaysia, Malaysia  
Prof. Varun Kumar Kakar, Kumaon Engineering College, Dwarahat, India  
Assist. Prof. Nisheeth Joshi, Apaji Institute, Banasthali University, Rajasthan, India  
Associate Prof. Kunwar S. Vaisla, VCT Kumaon Engineering College, India  
Prof Anupam Choudhary, Bhilai School Of Engg.,Bhilai (C.G.),India  
Mr. Divya Prakash Shrivastava, Al Jabal Al garbi University, Zawya, Libya

Associate Prof. Dr. V. Radha, Avinashilingam Deemed university for women, Coimbatore.  
Dr. Kasarapu Ramani, JNT University, Anantapur, India  
Dr. Anuraag Awasthi, Jayoti Vidyapeeth Womens University, India  
Dr. C G Ravichandran, R V S College of Engineering and Technology, India  
Dr. Mohamed A. Deriche, King Fahd University of Petroleum and Minerals, Saudi Arabia  
Mr. Abbas Karimi, Universiti Putra Malaysia, Malaysia  
Mr. Amit Kumar, Jaypee University of Engg. and Tech., India  
Dr. Nikolai Stoianov, Defense Institute, Bulgaria  
Assist. Prof. S. Ranichandra, KSR College of Arts and Science, Tiruchencode  
Mr. T.K.P. Rajagopal, Diamond Horse International Pvt Ltd, India  
Dr. Md. Ekramul Hamid, Rajshahi University, Bangladesh  
Mr. Hemanta Kumar Kalita, TATA Consultancy Services (TCS), India  
Dr. Messaouda Azzouzi, Ziane Achour University of Djelfa, Algeria  
Prof. (Dr.) Juan Jose Martinez Castillo, "Gran Mariscal de Ayacucho" University and Acantelys research Group, Venezuela  
Dr. Jatinderkumar R. Saini, Narmada College of Computer Application, India  
Dr. Babak Bashari Rad, University Technology of Malaysia, Malaysia  
Dr. Nighat Mir, Effat University, Saudi Arabia  
Prof. (Dr.) G.M.Nasira, Sasurie College of Engineering, India  
Mr. Varun Mittal, Gemalto Pte Ltd, Singapore  
Assist. Prof. Mrs P. Banumathi, Kathir College Of Engineering, Coimbatore  
Assist. Prof. Quan Yuan, University of Wisconsin-Stevens Point, US  
Dr. Pranam Paul, Narula Institute of Technology, Agarpara, West Bengal, India  
Assist. Prof. J. Ramkumar, V.L.B Janakiammal college of Arts & Science, India  
Mr. P. Sivakumar, Anna university, Chennai, India  
Mr. Md. Humayun Kabir Biswas, King Khalid University, Kingdom of Saudi Arabia  
Mr. Mayank Singh, J.P. Institute of Engg & Technology, Meerut, India  
HJ. Kamaruzaman Jusoff, Universiti Putra Malaysia  
Mr. Nikhil Patrick Lobo, CADES, India  
Dr. Amit Wason, Rayat-Bahra Institute of Engineering & Boi-Technology, India  
Dr. Rajesh Shrivastava, Govt. Benazir Science & Commerce College, Bhopal, India  
Assist. Prof. Vishal Bharti, DCE, Gurgaon  
Mrs. Sunita Bansal, Birla Institute of Technology & Science, India  
Dr. R. Sudhakar, Dr.Mahalingam college of Engineering and Technology, India  
Dr. Amit Kumar Garg, Shri Mata Vaishno Devi University, Katra(J&K), India  
Assist. Prof. Raj Gaurang Tiwari, AZAD Institute of Engineering and Technology, India  
Mr. Hamed Taherdoost, Tehran, Iran  
Mr. Amin Daneshmand Malayeri, YRC, IAU, Malayer Branch, Iran  
Mr. Shantanu Pal, University of Calcutta, India  
Dr. Terry H. Walcott, E-Promag Consultancy Group, United Kingdom  
Dr. Ezekiel U OKIKE, University of Ibadan, Nigeria  
Mr. P. Mahalingam, Caledonian College of Engineering, Oman

Dr. Mahmoud M. A. Abd Ellatif, Mansoura University, Egypt  
Prof. Kunwar S. Vaisla, BCT Kumaon Engineering College, India  
Prof. Mahesh H. Panchal, Kalol Institute of Technology & Research Centre, India  
Mr. Muhammad Asad, Technical University of Munich, Germany  
Mr. AliReza Shams Shafigh, Azad Islamic university, Iran  
Prof. S. V. Nagaraj, RMK Engineering College, India  
Mr. Ashikali M Hasan, Senior Researcher, CelNet security, India  
Dr. Adnan Shahid Khan, University Technology Malaysia, Malaysia  
Mr. Prakash Gajanan Burade, Nagpur University/ITM college of engg, Nagpur, India  
Dr. Jagdish B. Helonde, Nagpur University/ITM college of engg, Nagpur, India  
Professor, Doctor BOUHORMA Mohammed, University Abdelmalek Essaadi, Morocco  
Mr. K. Thirumalaivasan, Pondicherry Engg. College, India  
Mr. Umbarkar Anantkumar Janardan, Walchand College of Engineering, India  
Mr. Ashish Chaurasia, Gyan Ganga Institute of Technology & Sciences, India  
Mr. Sunil Taneja, Kurukshetra University, India  
Mr. Fauzi Adi Rafrastara, Dian Nuswantoro University, Indonesia  
Dr. Yaduvir Singh, Thapar University, India  
Dr. Ioannis V. Koskosas, University of Western Macedonia, Greece  
Dr. Vasantha Kalyani David, Avinashilingam University for women, Coimbatore  
Dr. Ahmed Mansour Manasrah, Universiti Sains Malaysia, Malaysia  
Miss. Nazanin Sadat Kazazi, University Technology Malaysia, Malaysia  
Mr. Saeed Rasouli Heikalabad, Islamic Azad University - Tabriz Branch, Iran  
Assoc. Prof. Dharendra Mishra, SVKM's NMIMS University, India  
Prof. Shapoor Zarei, UAE Inventors Association, UAE  
Prof. B.Raja Sarath Kumar, Lenora College of Engineering, India  
Dr. Bashir Alam, Jamia millia Islamia, Delhi, India  
Prof. Anant J Umbarkar, Walchand College of Engg., India  
Assist. Prof. B. Bharathi, Sathyabama University, India  
Dr. Fokrul Alom Mazarbhuiya, King Khalid University, Saudi Arabia  
Prof. T.S.Jeyali Laseeth, Anna University of Technology, Tirunelveli, India  
Dr. M. Balraju, Jawahar Lal Nehru Technological University Hyderabad, India  
Dr. Vijayalakshmi M. N., R.V.College of Engineering, Bangalore  
Prof. Walid Moudani, Lebanese University, Lebanon  
Dr. Saurabh Pal, VBS Purvanchal University, Jaunpur, India  
Associate Prof. Suneet Chaudhary, Dehradun Institute of Technology, India  
Associate Prof. Dr. Manuj Darbari, BBD University, India  
Ms. Prema Selvaraj, K.S.R College of Arts and Science, India  
Assist. Prof. Ms.S.Sasikala, KSR College of Arts & Science, India  
Mr. Sukhvinder Singh Deora, NC Institute of Computer Sciences, India  
Dr. Abhay Bansal, Amity School of Engineering & Technology, India  
Ms. Sumita Mishra, Amity School of Engineering and Technology, India  
Professor S. Viswanadha Raju, JNT University Hyderabad, India

Mr. Asghar Shahrzad Khashandarag, Islamic Azad University Tabriz Branch, India  
Mr. Manoj Sharma, Panipat Institute of Engg. & Technology, India  
Mr. Shakeel Ahmed, King Faisal University, Saudi Arabia  
Dr. Mohamed Ali Mahjoub, Institute of Engineer of Monastir, Tunisia  
Mr. Adri Jovin J.J., SriGuru Institute of Technology, India  
Dr. Sukumar Senthilkumar, Universiti Sains Malaysia, Malaysia  
Mr. Rakesh Bharati, Dehradun Institute of Technology Dehradun, India  
Mr. Shervan Fekri Ershad, Shiraz International University, Iran  
Mr. Md. Safiqul Islam, Daffodil International University, Bangladesh  
Mr. Mahmudul Hasan, Daffodil International University, Bangladesh  
Prof. Mandakini Tayade, UIT, RGTU, Bhopal, India  
Ms. Sarla More, UIT, RGTU, Bhopal, India  
Mr. Tushar Hrishikesh Jaware, R.C. Patel Institute of Technology, Shirpur, India  
Ms. C. Divya, Dr G R Damodaran College of Science, Coimbatore, India  
Mr. Fahimuddin Shaik, Annamacharya Institute of Technology & Sciences, India  
Dr. M. N. Giri Prasad, JNTUCE, Pulivendula, A.P., India  
Assist. Prof. Chintan M Bhatt, Charotar University of Science And Technology, India  
Prof. Sahista Machchhar, Marwadi Education Foundation's Group of institutions, India  
Assist. Prof. Navnish Goel, S. D. College Of Enginnering & Technology, India  
Mr. Khaja Kamaluddin, Sirt University, Sirt, Libya  
Mr. Mohammad Zaidul Karim, Daffodil International, Bangladesh  
Mr. M. Vijayakumar, KSR College of Engineering, Tiruchengode, India  
Mr. S. A. Ahsan Rajon, Khulna University, Bangladesh  
Dr. Muhammad Mohsin Nazir, LCW University Lahore, Pakistan  
Mr. Mohammad Asadul Hoque, University of Alabama, USA  
Mr. P.V.Sarathchand, Indur Institute of Engineering and Technology, India  
Mr. Durgesh Samadhiya, Chung Hua University, Taiwan  
Dr Venu Kuthadi, University of Johannesburg, Johannesburg, RSA  
Dr. (Er) Jasvir Singh, Guru Nanak Dev University, Amritsar, Punjab, India  
Mr. Jasmin Cosic, Min. of the Interior of Una-sana canton, B&H, Bosnia and Herzegovina  
Dr S. Rajalakshmi, Botho College, South Africa  
Dr. Mohamed Sarrab, De Montfort University, UK  
Mr. Basappa B. Kodada, Canara Engineering College, India  
Assist. Prof. K. Ramana, Annamacharya Institute of Technology and Sciences, India  
Dr. Ashu Gupta, Apeejay Institute of Management, Jalandhar, India  
Assist. Prof. Shaik Rasool, Shadan College of Engineering & Technology, India  
Assist. Prof. K. Suresh, Annamacharya Institute of Tech & Sci. Rajampet, AP, India  
Dr . G. Singaravel, K.S.R. College of Engineering, India  
Dr B. G. Geetha, K.S.R. College of Engineering, India  
Assist. Prof. Kavita Choudhary, ITM University, Gurgaon  
Dr. Mehrdad Jalali, Azad University, Mashhad, Iran  
Megha Goel, Shamli Institute of Engineering and Technology, Shamli, India

Mr. Chi-Hua Chen, Institute of Information Management, National Chiao-Tung University, Taiwan (R.O.C.)  
Assoc. Prof. A. Rajendran, RVS College of Engineering and Technology, India  
Assist. Prof. S. Jaganathan, RVS College of Engineering and Technology, India  
Assoc. Prof. (Dr.) A S N Chakravarthy, JNTUK University College of Engineering Vizianagaram (State University)  
Assist. Prof. Deepshikha Patel, Technocrat Institute of Technology, India  
Assist. Prof. Maram Balajee, GMRIT, India  
Assist. Prof. Monika Bhatnagar, TIT, India  
Prof. Gaurang Panchal, Charotar University of Science & Technology, India  
Prof. Anand K. Tripathi, Computer Society of India  
Prof. Jyoti Chaudhary, High Performance Computing Research Lab, India  
Assist. Prof. Supriya Raheja, ITM University, India  
Dr. Pankaj Gupta, Microsoft Corporation, U.S.A.  
Assist. Prof. Panchamukesh Chandaka, Hyderabad Institute of Tech. & Management, India  
Prof. Mohan H.S, SJB Institute Of Technology, India  
Mr. Hossein Malekinezhad, Islamic Azad University, Iran  
Mr. Zatin Gupta, Universti Malaysia, Malaysia  
Assist. Prof. Amit Chauhan, Phonics Group of Institutions, India  
Assist. Prof. Ajal A. J., METS School Of Engineering, India  
Mrs. Omowunmi Omobola Adeyemo, University of Ibadan, Nigeria  
Dr. Bharat Bhushan Agarwal, I.F.T.M. University, India  
Md. Nazrul Islam, University of Western Ontario, Canada  
Tushar Kanti, L.N.C.T, Bhopal, India  
Er. Aumreesh Kumar Saxena, SIRTs College Bhopal, India  
Mr. Mohammad Monirul Islam, Daffodil International University, Bangladesh  
Dr. Kashif Nisar, University Utara Malaysia, Malaysia  
Dr. Wei Zheng, Rutgers Univ/ A10 Networks, USA  
Associate Prof. Rituraj Jain, Vyas Institute of Engg & Tech, Jodhpur – Rajasthan  
Assist. Prof. Apoorvi Sood, I.T.M. University, India  
Dr. Kayhan Zrar Ghafoor, University Technology Malaysia, Malaysia  
Mr. Swapnil Sonar, Truba Institute College of Engineering & Technology, Indore, India  
Ms. Yogita Gigras, I.T.M. University, India  
Associate Prof. Neelima Sadineni, Pydha Engineering College, India Pydha Engineering College  
Assist. Prof. K. Deepika Rani, HITAM, Hyderabad  
Ms. Shikha Maheshwari, Jaipur Engineering College & Research Centre, India  
Prof. Dr V S Giridhar Akula, Avanthi's Scientific Tech. & Research Academy, Hyderabad  
Prof. Dr.S.Saravanan, Muthayammal Engineering College, India  
Mr. Mehdi Golsorkhatabar Amiri, Islamic Azad University, Iran  
Prof. Amit Sadanand Savyanavar, MITCOE, Pune, India  
Assist. Prof. P.Oliver Jayaprakash, Anna University, Chennai  
Assist. Prof. Ms. Sujata, ITM University, Gurgaon, India  
Dr. Asoke Nath, St. Xavier's College, India

Mr. Masoud Rafighi, Islamic Azad University, Iran  
Assist. Prof. RamBabu Pemula, NIMRA College of Engineering & Technology, India  
Assist. Prof. Ms Rita Chhikara, ITM University, Gurgaon, India  
Mr. Sandeep Maan, Government Post Graduate College, India  
Prof. Dr. S. Muralidharan, Mepco Schlenk Engineering College, India  
Associate Prof. T.V.Sai Krishna, QIS College of Engineering and Technology, India  
Mr. R. Balu, Bharathiar University, Coimbatore, India  
Assist. Prof. Shekhar. R, Dr.SM College of Engineering, India  
Prof. P. Senthilkumar, Vivekanandha Institute of Engineering and Technology for Woman, India  
Mr. M. Kamarajan, PSNA College of Engineering & Technology, India  
Dr. Angajala Srinivasa Rao, Jawaharlal Nehru Technical University, India  
Assist. Prof. C. Venkatesh, A.I.T.S, Rajampet, India  
Mr. Afshin Rezakhani Roozbahani, Ayatollah Boroujerdi University, Iran  
Mr. Laxmi chand, SCTL, Noida, India  
Dr. Dr. Abdul Hannan, Vivekanand College, Aurangabad  
Prof. Mahesh Panchal, KITRC, Gujarat  
Dr. A. Subramani, K.S.R. College of Engineering, Tiruchengode  
Assist. Prof. Prakash M, Rajalakshmi Engineering College, Chennai, India  
Assist. Prof. Akhilesh K Sharma, Sir Padampat Singhania University, India  
Ms. Varsha Sahni, Guru Nanak Dev Engineering College, Ludhiana, India  
Associate Prof. Trilochan Rout, NM Institute of Engineering and Technology, India  
Mr. Srikantha Kumar Mohapatra, NMIET, Orissa, India  
Mr. Waqas Haider Bangyal, Iqra University Islamabad, Pakistan  
Dr. S. Vijayaragavan, Christ College of Engineering and Technology, Pondicherry, India  
Prof. Elboukhari Mohamed, University Mohammed First, Oujda, Morocco  
Dr. Muhammad Asif Khan, King Faisal University, Saudi Arabia  
Dr. Nagy Ramadan Darwish Omran, Cairo University, Egypt.  
Assistant Prof. Anand Nayyar, KCL Institute of Management and Technology, India  
Mr. G. Premsankar, Ericsson, India  
Assist. Prof. T. Hemalatha, VELS University, India  
Prof. Tejaswini Apte, University of Pune, India  
Dr. Edmund Ng Giap Weng, Universiti Malaysia Sarawak, Malaysia  
Mr. Mahdi Nouri, Iran University of Science and Technology, Iran  
Associate Prof. S. Asif Hussain, Annamacharya Institute of technology & Sciences, India  
Mrs. Kavita Pabreja, Maharaja Surajmal Institute (an affiliate of GGSIP University), India  
Mr. Vorugunti Chandra Sekhar, DA-IICT, India  
Mr. Muhammad Najmi Ahmad Zabidi, Universiti Teknologi Malaysia, Malaysia  
Dr. Aderemi A. Atayero, Covenant University, Nigeria  
Assist. Prof. Osama Sohaib, Balochistan University of Information Technology, Pakistan  
Assist. Prof. K. Suresh, Annamacharya Institute of Technology and Sciences, India  
Mr. Hassen Mohammed Abdullaah Alsafi, International Islamic University Malaysia (IIUM) Malaysia  
Mr. Robail Yasrab, Virtual University of Pakistan, Pakistan



Mr. R. Balu, Bharathiar University, Coimbatore, India  
Prof. Anand Nayyar, KCL Institute of Management and Technology, Jalandhar  
Assoc. Prof. Vivek S Deshpande, MIT College of Engineering, India  
Prof. K. Saravanan, Anna university Coimbatore, India  
Dr. Ravendra Singh, MJP Rohilkhand University, Bareilly, India  
Mr. V. Mathivanan, IBRA College of Technology, Sultanate of OMAN  
Assoc. Prof. S. Asif Hussain, AITS, India  
Assist. Prof. C. Venkatesh, AITS, India  
Mr. Sami Ulhaq, SZABIST Islamabad, Pakistan  
Dr. B. Justus Rabi, Institute of Science & Technology, India  
Mr. Anuj Kumar Yadav, Dehradun Institute of technology, India  
Mr. Alejandro Mosquera, University of Alicante, Spain  
Assist. Prof. Arjun Singh, Sir Padampat Singhania University (SPSU), Udaipur, India  
Dr. Smriti Agrawal, JB Institute of Engineering and Technology, Hyderabad  
Assist. Prof. Swathi Sambangi, Visakha Institute of Engineering and Technology, India  
Ms. Prabhjot Kaur, Guru Gobind Singh Indraprastha University, India  
Mrs. Samaher AL-Hothali, Yanbu University College, Saudi Arabia  
Prof. Rajneeshkaur Bedi, MIT College of Engineering, Pune, India  
Mr. Hassen Mohammed Abdullah Alsafi, International Islamic University Malaysia (IIUM)  
Dr. Wei Zhang, Amazon.com, Seattle, WA, USA  
Mr. B. Santhosh Kumar, C S I College of Engineering, Tamil Nadu  
Dr. K. Reji Kumar, , N S S College, Pandalam, India  
Assoc. Prof. K. Seshadri Sastry, EIILM University, India  
Mr. Kai Pan, UNC Charlotte, USA  
Mr. Ruikar Sachin, SGGSIET, India  
Prof. (Dr.) Vinodani Katiyar, Sri Ramswaroop Memorial University, India  
Assoc. Prof., M. Giri, Sreenivasa Institute of Technology and Management Studies, India  
Assoc. Prof. Labib Francis Gergis, Misr Academy for Engineering and Technology (MET), Egypt  
Assist. Prof. Amanpreet Kaur, ITM University, India  
Assist. Prof. Anand Singh Rajawat, Shri Vaishnav Institute of Technology & Science, Indore  
Mrs. Hadeel Saleh Haj Aliwi, Universiti Sains Malaysia (USM), Malaysia  
Dr. Abhay Bansal, Amity University, India  
Dr. Mohammad A. Mezher, Fahad Bin Sultan University, KSA  
Assist. Prof. Nidhi Arora, M.C.A. Institute, India  
Prof. Dr. P. Suresh, Karpagam College of Engineering, Coimbatore, India  
Dr. Kannan Balasubramanian, Mepco Schlenk Engineering College, India  
Dr. S. Sankara Gomathi, Panimalar Engineering college, India  
Prof. Anil kumar Suthar, Gujarat Technological University, L.C. Institute of Technology, India  
Assist. Prof. R. Hubert Rajan, NOORUL ISLAM UNIVERSITY, India  
Assist. Prof. Dr. Jyoti Mahajan, College of Engineering & Technology  
Assist. Prof. Homam Reda El-Taj, College of Network Engineering, Saudi Arabia & Malaysia  
Mr. Bijan Paul, Shahjalal University of Science & Technology, Bangladesh

Assoc. Prof. Dr. Ch V Phani Krishna, KL University, India  
Dr. Vishal Bhatnagar, Ambedkar Institute of Advanced Communication Technologies & Research, India  
Dr. Lamri LAOUAMER, Al Qassim University, Dept. Info. Systems & European University of Brittany, Dept.  
Computer Science, UBO, Brest, France  
Prof. Ashish Babanrao Sasankar, G.H.Raisoni Institute Of Information Technology, India  
Prof. Pawan Kumar Goel, Shamli Institute of Engineering and Technology, India  
Mr. Ram Kumar Singh, S.V Subharti University, India  
Assistant Prof. Sunish Kumar O S, Amaljiyothi College of Engineering, India  
Dr Sanjay Bhargava, Banasthali University, India  
Mr. Pankaj S. Kulkarni, AVEW's Shatabdi Institute of Technology, India  
Mr. Roohollah Etemadi, Islamic Azad University, Iran  
Mr. Oloruntoyin Sefiu Taiwo, Emmanuel Alayande College Of Education, Nigeria  
Mr. Sumit Goyal, National Dairy Research Institute, India  
Mr Jaswinder Singh Dilawari, Geeta Engineering College, India  
Prof. Raghuraj Singh, Harcourt Butler Technological Institute, Kanpur  
Dr. S.K. Mahendran, Anna University, Chennai, India  
Dr. Amit Wason, Hindustan Institute of Technology & Management, Punjab  
Dr. Ashu Gupta, Apeejay Institute of Management, India  
Assist. Prof. D. Asir Antony Gnana Singh, M.I.E.T Engineering College, India  
Mrs Mina Farmanbar, Eastern Mediterranean University, Famagusta, North Cyprus  
Mr. Maram Balajee, GMR Institute of Technology, India  
Mr. Moiz S. Ansari, Isra University, Hyderabad, Pakistan  
Mr. Adebayo, Olawale Surajudeen, Federal University of Technology Minna, Nigeria  
Mr. Jasvir Singh, University College Of Engg., India  
Mr. Vivek Tiwari, MANIT, Bhopal, India  
Assoc. Prof. R. Navaneethakrishnan, Bharathiyar College of Engineering and Technology, India  
Mr. Somdip Dey, St. Xavier's College, Kolkata, India  
Mr. Souleymane Balla-Arabé, Xi'an University of Electronic Science and Technology, China  
Mr. Mahabub Alam, Rajshahi University of Engineering and Technology, Bangladesh  
Mr. Sathyapraksh P., S.K.P Engineering College, India  
Dr. N. Karthikeyan, SNS College of Engineering, Anna University, India  
Dr. Binod Kumar, JSPM's, Jayawant Technical Campus, Pune, India  
Assoc. Prof. Dinesh Goyal, Suresh Gyan Vihar University, India  
Mr. Md. Abdul Ahad, K L University, India  
Mr. Vikas Bajpai, The LNM IIT, India  
Dr. Manish Kumar Anand, Salesforce (R & D Analytics), San Francisco, USA  
Assist. Prof. Dheeraj Murari, Kumaon Engineering College, India  
Assoc. Prof. Dr. A. Muthukumaravel, VELS University, Chennai  
Mr. A. Siles Balasingh, St.Joseph University in Tanzania, Tanzania  
Mr. Ravindra Daga Badgujar, R C Patel Institute of Technology, India  
Dr. Preeti Khanna, SVKM's NMIMS, School of Business Management, India  
Mr. Kumar Dayanand, Cambridge Institute of Technology, India

Dr. Syed Asif Ali, SMI University Karachi, Pakistan  
Prof. Pallvi Pandit, Himachal Pradesh University, India  
Mr. Ricardo Verschuere, University of Gloucestershire, UK  
Assist. Prof. Mamta Juneja, University Institute of Engineering and Technology, Panjab University, India  
Assoc. Prof. P. Surendra Varma, NRI Institute of Technology, JNTU Kakinada, India  
Assist. Prof. Gaurav Shrivastava, RGPV / SVITS Indore, India  
Dr. S. Sumathi, Anna University, India  
Assist. Prof. Ankita M. Kapadia, Charotar University of Science and Technology, India  
Mr. Deepak Kumar, Indian Institute of Technology (BHU), India  
Dr. Dr. Rajan Gupta, GGSIP University, New Delhi, India  
Assist. Prof. M. Anand Kumar, Karpagam University, Coimbatore, India  
Mr. Arshad Mansoor, Pakistan Aeronautical Complex  
Mr. Kapil Kumar Gupta, Ansal Institute of Technology and Management, India  
Dr. Neeraj Tomer, SINE International Institute of Technology, Jaipur, India  
Assist. Prof. Trunal J. Patel, C.G. Patel Institute of Technology, Uka Tarsadia University, Bardoli, Surat  
Mr. Sivakumar, Codework solutions, India  
Mr. Mohammad Sadegh Mirzaei, PGNR Company, Iran  
Dr. Gerard G. Dumancas, Oklahoma Medical Research Foundation, USA  
Mr. Varadala Sridhar, Varadhaman College Engineering College, Affiliated To JNTU, Hyderabad  
Assist. Prof. Manoj Dhawan, SVITS, Indore  
Assoc. Prof. Chitresh Banerjee, Suresh Gyan Vihar University, Jaipur, India  
Dr. S. Santhi, SCSVMV University, India  
Mr. Davood Mohammadi Souran, Ministry of Energy of Iran, Iran  
Mr. Shamim Ahmed, Bangladesh University of Business and Technology, Bangladesh  
Mr. Sandeep Reddivari, Mississippi State University, USA  
Assoc. Prof. Ousmane Thiare, Gaston Berger University, Senegal  
Dr. Hazra Imran, Athabasca University, Canada  
Dr. Setu Kumar Chaturvedi, Technocrats Institute of Technology, Bhopal, India  
Mr. Mohd Dilshad Ansari, Jaypee University of Information Technology, India  
Ms. Jaspreet Kaur, Distance Education LPU, India  
Dr. D. Nagarajan, Salalah College of Technology, Sultanate of Oman  
Dr. K.V.N.R.Sai Krishna, S.V.R.M. College, India  
Mr. Himanshu Pareek, Center for Development of Advanced Computing (CDAC), India  
Mr. Khaldi Amine, Badji Mokhtar University, Algeria  
Mr. Mohammad Sadegh Mirzaei, Scientific Applied University, Iran  
Assist. Prof. Khyati Chaudhary, Ram-eesh Institute of Engg. & Technology, India  
Mr. Sanjay Agal, Pacific College of Engineering Udaipur, India  
Mr. Abdul Mateen Ansari, King Khalid University, Saudi Arabia  
Dr. H.S. Behera, Veer Surendra Sai University of Technology (VSSUT), India  
Dr. Shrikant Tiwari, Shri Shankaracharya Group of Institutions (SSGI), India  
Prof. Ganesh B. Regulwar, Shri Shankarprasad Agnihotri College of Engg, India  
Prof. Pinmananeni Bhanu Prasad, Matrix vision GmbH, Germany

Dr. Shrikant Tiwari, Shri Shankaracharya Technical Campus (SSTC), India  
Dr. Siddesh G.K., : Dayananada Sagar College of Engineering, Bangalore, India  
Dr. Nadir Bouchama, CERIST Research Center, Algeria  
Dr. R. Sathishkumar, Sri Venkateswara College of Engineering, India  
Assistant Prof (Dr.) Mohamed Moussaoui, Abdelmalek Essaadi University, Morocco  
Dr. S. Malathi, Panimalar Engineering College, Chennai, India  
Dr. V. Subedha, Panimalar Institute of Technology, Chennai, India  
Dr. Prashant Panse, Swami Vivekanand College of Engineering, Indore, India  
Dr. Hamza Aldabbas, Al-Balqa'a Applied University, Jordan  
Dr. G. Rasitha Banu, Vel's University, Chennai  
Dr. V. D. Ambeth Kumar, Panimalar Engineering College, Chennai  
Prof. Anuranjan Misra, Bhagwant Institute of Technology, Ghaziabad, India  
Ms. U. Sinthuja, PSG college of arts & science, India  
Dr. Ehsan Saradar Torshizi, Urmia University, Iran  
Dr. Shamneesh Sharma, APG Shimla University, Shimla (H.P.), India  
Assistant Prof. A. S. Syed Navaz, Muthayammal College of Arts & Science, India  
Assistant Prof. Ranjit Panigrahi, Sikkim Manipal Institute of Technology, Majitar, Sikkim  
Dr. Khaled Eskaf, Arab Academy for Science ,Technology & Maritime Transportation, Egypt  
Dr. Nishant Gupta, University of Jammu, India  
Assistant Prof. Nagarajan Sankaran, Annamalai University, Chidambaram, Tamilnadu, India  
Assistant Prof. Tribikram Pradhan, Manipal Institute of Technology, India  
Dr. Nasser Lotfi, Eastern Mediterranean University, Northern Cyprus  
Dr. R. Manavalan, K S Rangasamy college of Arts and Science, Tamilnadu, India  
Assistant Prof. P. Krishna Sankar, K S Rangasamy college of Arts and Science, Tamilnadu, India  
Dr. Rahul Malik, Cisco Systems, USA  
Dr. S. C. Lingareddy, ALPHA College of Engineering, India  
Assistant Prof. Mohammed Shuaib, Interl University, Lucknow, India  
Dr. Sachin Yele, Sanghvi Institute of Management & Science, India  
Dr. T. Thambidurai, Sun Univercell, Singapore  
Prof. Anandkumar Telang, BKIT, India  
Assistant Prof. R. Poorvadevi, SCSVMV University, India  
Dr Uttam Mande, Gitam University, India  
Dr. Poornima Girish Naik, Shahu Institute of Business Education and Research (SIBER), India  
Prof. Md. Abu Kausar, Jaipur National University, Jaipur, India  
Dr. Mohammed Zuber, AISECT University, India  
Prof. Kalum Priyanath Udagepola, King Abdulaziz University, Saudi Arabia  
Dr. K. R. Ananth, Velalar College of Engineering and Technology, India  
Assistant Prof. Sanjay Sharma, Roorkee Engineering & Management Institute Shamli (U.P), India  
Assistant Prof. Panem Charan Arur, Priyadarshini Institute of Technology, India  
Dr. Ashwak Mahmood muhsen alabaichi, Karbala University / College of Science, Iraq  
Dr. Urmila Shrawankar, G H Raison College of Engineering, Nagpur (MS), India  
Dr. Krishan Kumar Paliwal, Panipat Institute of Engineering & Technology, India

Dr. Mukesh Negi, Tech Mahindra, India  
Dr. Anuj Kumar Singh, Amity University Gurgaon, India  
Dr. Babar Shah, Gyeongsang National University, South Korea  
Assistant Prof. Jayprakash Upadhyay, SRI-TECH Jabalpur, India  
Assistant Prof. Varadala Sridhar, Vidya Jyothi Institute of Technology, India  
Assistant Prof. Parameshachari B D, KSIT, Bangalore, India  
Assistant Prof. Ankit Garg, Amity University, Haryana, India  
Assistant Prof. Rajashe Karappa, SDMCET, Karnataka, India  
Assistant Prof. Varun Jasuja, GNIT, India  
Assistant Prof. Sonal Honale, Abha Gaikwad Patil College of Engineering Nagpur, India  
Dr. Pooja Choudhary, CT Group of Institutions, NIT Jalandhar, India  
Dr. Faouzi Hidoussi, UHL Batna, Algeria  
Dr. Naseer Ali Husieen, Wasit University, Iraq  
Assistant Prof. Vinod Kumar Shukla, Amity University, Dubai  
Dr. Ahmed Farouk Metwaly, K L University  
Mr. Mohammed Noaman Murad, Cihan University, Iraq  
Dr. Suxing Liu, Arkansas State University, USA  
Dr. M. Gomathi, Velalar College of Engineering and Technology, India

# **CALL FOR PAPERS**

## **International Journal of Computer Science and Information Security**

**IJCSIS 2014**

**ISSN: 1947-5500**

**<http://sites.google.com/site/ijcsis/>**

International Journal Computer Science and Information Security, IJCSIS, is the premier scholarly venue in the areas of computer science and security issues. IJCSIS 2011 will provide a high profile, leading edge platform for researchers and engineers alike to publish state-of-the-art research in the respective fields of information technology and communication security. The journal will feature a diverse mixture of publication articles including core and applied computer science related topics.

Authors are solicited to contribute to the special issue by submitting articles that illustrate research results, projects, surveying works and industrial experiences that describe significant advances in the following areas, but are not limited to. Submissions may span a broad range of topics, e.g.:

### ***Track A: Security***

Access control, Anonymity, Audit and audit reduction & Authentication and authorization, Applied cryptography, Cryptanalysis, Digital Signatures, Biometric security, Boundary control devices, Certification and accreditation, Cross-layer design for security, Security & Network Management, Data and system integrity, Database security, Defensive information warfare, Denial of service protection, Intrusion Detection, Anti-malware, Distributed systems security, Electronic commerce, E-mail security, Spam, Phishing, E-mail fraud, Virus, worms, Trojan Protection, Grid security, Information hiding and watermarking & Information survivability, Insider threat protection, Integrity

Intellectual property protection, Internet/Intranet Security, Key management and key recovery, Language-based security, Mobile and wireless security, Mobile, Ad Hoc and Sensor Network Security, Monitoring and surveillance, Multimedia security ,Operating system security, Peer-to-peer security, Performance Evaluations of Protocols & Security Application, Privacy and data protection, Product evaluation criteria and compliance, Risk evaluation and security certification, Risk/vulnerability assessment, Security & Network Management, Security Models & protocols, Security threats & countermeasures (DDoS, MiM, Session Hijacking, Replay attack etc.), Trusted computing, Ubiquitous Computing Security, Virtualization security, VoIP security, Web 2.0 security, Submission Procedures, Active Defense Systems, Adaptive Defense Systems, Benchmark, Analysis and Evaluation of Security Systems, Distributed Access Control and Trust Management, Distributed Attack Systems and Mechanisms, Distributed Intrusion Detection/Prevention Systems, Denial-of-Service Attacks and Countermeasures, High Performance Security Systems, Identity Management and Authentication, Implementation, Deployment and Management of Security Systems, Intelligent Defense Systems, Internet and Network Forensics, Large-scale Attacks and Defense, RFID Security and Privacy, Security Architectures in Distributed Network Systems, Security for Critical Infrastructures, Security for P2P systems and Grid Systems, Security in E-Commerce, Security and Privacy in Wireless Networks, Secure Mobile Agents and Mobile Code, Security Protocols, Security Simulation and Tools, Security Theory and Tools, Standards and Assurance Methods, Trusted Computing, Viruses, Worms, and Other Malicious Code, World Wide Web Security, Novel and emerging secure architecture, Study of attack strategies, attack modeling, Case studies and analysis of actual attacks, Continuity of Operations during an attack, Key management, Trust management, Intrusion detection techniques, Intrusion response, alarm management, and correlation analysis, Study of tradeoffs between security and system performance, Intrusion tolerance systems, Secure protocols, Security in wireless networks (e.g. mesh networks, sensor networks, etc.), Cryptography and Secure Communications, Computer Forensics, Recovery and Healing, Security Visualization, Formal Methods in Security, Principles for Designing a Secure Computing System, Autonomic Security, Internet Security, Security in Health Care Systems, Security Solutions Using Reconfigurable Computing, Adaptive and Intelligent Defense Systems, Authentication and Access control, Denial of service attacks and countermeasures, Identity, Route and

Location Anonymity schemes, Intrusion detection and prevention techniques, Cryptography, encryption algorithms and Key management schemes, Secure routing schemes, Secure neighbor discovery and localization, Trust establishment and maintenance, Confidentiality and data integrity, Security architectures, deployments and solutions, Emerging threats to cloud-based services, Security model for new services, Cloud-aware web service security, Information hiding in Cloud Computing, Securing distributed data storage in cloud, Security, privacy and trust in mobile computing systems and applications, **Middleware security & Security features:** middleware software is an asset on

its own and has to be protected, interaction between security-specific and other middleware features, e.g., context-awareness, **Middleware-level security monitoring and measurement:** metrics and mechanisms for quantification and evaluation of security enforced by the middleware, **Security co-design:** trade-off and co-design between application-based and middleware-based security, **Policy-based management:** innovative support for policy-based definition and enforcement of security concerns, **Identification and authentication mechanisms:** Means to capture application specific constraints in defining and enforcing access control rules, **Middleware-oriented security patterns:** identification of patterns for sound, reusable security, **Security in aspect-based middleware:** mechanisms for isolating and enforcing security aspects, **Security in agent-based platforms:** protection for mobile code and platforms, Smart Devices: Biometrics, National ID cards, Embedded Systems Security and TPMs, RFID Systems Security, Smart Card Security, Pervasive Systems: Digital Rights Management (DRM) in pervasive environments, Intrusion Detection and Information Filtering, Localization Systems Security (Tracking of People and Goods), Mobile Commerce Security, Privacy Enhancing Technologies, Security Protocols (for Identification and Authentication, Confidentiality and Privacy, and Integrity), Ubiquitous Networks: Ad Hoc Networks Security, Delay-Tolerant Network Security, Domestic Network Security, Peer-to-Peer Networks Security, Security Issues in Mobile and Ubiquitous Networks, Security of GSM/GPRS/UMTS Systems, Sensor Networks Security, Vehicular Network Security, Wireless Communication Security: Bluetooth, NFC, WiFi, WiMAX, WiMedia, others

This Track will emphasize the design, implementation, management and applications of computer communications, networks and services. Topics of mostly theoretical nature are also welcome, provided there is clear practical potential in applying the results of such work.

### ***Track B: Computer Science***

Broadband wireless technologies: LTE, WiMAX, WiRAN, HSDPA, HSUPA, Resource allocation and interference management, Quality of service and scheduling methods, Capacity planning and dimensioning, Cross-layer design and Physical layer based issue, Interworking architecture and interoperability, Relay assisted and cooperative communications, Location and provisioning and mobility management, Call admission and flow/congestion control, Performance optimization, Channel capacity modeling and analysis, Middleware Issues: Event-based, publish/subscribe, and message-oriented middleware, Reconfigurable, adaptable, and reflective middleware approaches, Middleware solutions for reliability, fault tolerance, and quality-of-service, Scalability of middleware, Context-aware middleware, Autonomic and self-managing middleware, Evaluation techniques for middleware solutions, Formal methods and tools for designing, verifying, and evaluating, middleware, Software engineering techniques for middleware, Service oriented middleware, Agent-based middleware, Security middleware, Network Applications: Network-based automation, Cloud applications, Ubiquitous and pervasive applications, Collaborative applications, RFID and sensor network applications, Mobile applications, Smart home applications, Infrastructure monitoring and control applications, Remote health monitoring, GPS and location-based applications, Networked vehicles applications, Alert applications, Embedded Computer System, Advanced Control Systems, and Intelligent Control : Advanced control and measurement, computer and microprocessor-based control, signal processing, estimation and identification techniques, application specific IC's, nonlinear and adaptive control, optimal and robot control, intelligent control, evolutionary computing, and intelligent systems, instrumentation subject to critical conditions, automotive, marine and aero-space control and all other control applications, Intelligent Control System, Wiring/Wireless Sensor, Signal Control System. Sensors, Actuators and Systems Integration : Intelligent sensors and actuators, multisensor fusion, sensor array and multi-channel processing, micro/nano technology, microsensors and microactuators, instrumentation electronics, MEMS and system integration, wireless sensor, Network Sensor, Hybrid



Sensor, Distributed Sensor Networks. Signal and Image Processing : Digital signal processing theory, methods, DSP implementation, speech processing, image and multidimensional signal processing, Image analysis and processing, Image and Multimedia applications, Real-time multimedia signal processing, Computer vision, Emerging signal processing areas, Remote Sensing, Signal processing in education. Industrial Informatics: Industrial applications of neural networks, fuzzy algorithms, Neuro-Fuzzy application, bioInformatics, real-time computer control, real-time information systems, human-machine interfaces, CAD/CAM/CAT/CIM, virtual reality, industrial communications, flexible manufacturing systems, industrial automated process, Data Storage Management, Harddisk control, Supply Chain Management, Logistics applications, Power plant automation, Drives automation. Information Technology, Management of Information System : Management information systems, Information Management, Nursing information management, Information System, Information Technology and their application, Data retrieval, Data Base Management, Decision analysis methods, Information processing, Operations research, E-Business, E-Commerce, E-Government, Computer Business, Security and risk management, Medical imaging, Biotechnology, Bio-Medicine, Computer-based information systems in health care, Changing Access to Patient Information, Healthcare Management Information Technology. Communication/Computer Network, Transportation Application : On-board diagnostics, Active safety systems, Communication systems, Wireless technology, Communication application, Navigation and Guidance, Vision-based applications, Speech interface, Sensor fusion, Networking theory and technologies, Transportation information, Autonomous vehicle, Vehicle application of affective computing, Advance Computing technology and their application : Broadband and intelligent networks, Data Mining, Data fusion, Computational intelligence, Information and data security, Information indexing and retrieval, Information processing, Information systems and applications, Internet applications and performances, Knowledge based systems, Knowledge management, Software Engineering, Decision making, Mobile networks and services, Network management and services, Neural Network, Fuzzy logics, Neuro-Fuzzy, Expert approaches, Innovation Technology and Management : Innovation and product development, Emerging advances in business and its applications, Creativity in Internet management and retailing, B2B and B2C management, Electronic transceiver device for Retail Marketing Industries, Facilities planning and management, Innovative pervasive computing applications, Programming paradigms for pervasive systems, Software evolution and maintenance in pervasive systems, Middleware services and agent technologies, Adaptive, autonomic and context-aware computing, Mobile/Wireless computing systems and services in pervasive computing, Energy-efficient and green pervasive computing, Communication architectures for pervasive computing, Ad hoc networks for pervasive communications, Pervasive opportunistic communications and applications, Enabling technologies for pervasive systems (e.g., wireless BAN, PAN), Positioning and tracking technologies, Sensors and RFID in pervasive systems, Multimodal sensing and context for pervasive applications, Pervasive sensing, perception and semantic interpretation, Smart devices and intelligent environments, Trust, security and privacy issues in pervasive systems, User interfaces and interaction models, Virtual immersive communications, Wearable computers, Standards and interfaces for pervasive computing environments, Social and economic models for pervasive systems, Active and Programmable Networks, Ad Hoc & Sensor Network, Congestion and/or Flow Control, Content Distribution, Grid Networking, High-speed Network Architectures, Internet Services and Applications, Optical Networks, Mobile and Wireless Networks, Network Modeling and Simulation, Multicast, Multimedia Communications, Network Control and Management, Network Protocols, Network Performance, Network Measurement, Peer to Peer and Overlay Networks, Quality of Service and Quality of Experience, Ubiquitous Networks, Crosscutting Themes – Internet Technologies, Infrastructure, Services and Applications; Open Source Tools, Open Models and Architectures; Security, Privacy and Trust; Navigation Systems, Location Based Services; Social Networks and Online Communities; ICT Convergence, Digital Economy and Digital Divide, Neural Networks, Pattern Recognition, Computer Vision, Advanced Computing Architectures and New Programming Models, Visualization and Virtual Reality as Applied to Computational Science, Computer Architecture and Embedded Systems, Technology in Education, Theoretical Computer Science, Computing Ethics, Computing Practices & Applications

Authors are invited to submit papers through e-mail [ijcsiseditor@gmail.com](mailto:ijcsiseditor@gmail.com). Submissions must be original and should not have been published previously or be under consideration for publication while being evaluated by IJCSIS. Before submission authors should carefully read over the journal's Author Guidelines, which are located at <http://sites.google.com/site/ijcsis/authors-notes> .



**© IJCSIS PUBLICATION 2014**  
**ISSN 1947 5500**  
**<http://sites.google.com/site/ijcsis/>**